Profiling of Oligolignols Reveals Monolignol Coupling Conditions in Lignifying Poplar Xylem[1][w]

Kris Morreel, John Ralph, Hoon Kim, Fachuang Lu, Geert Goeminne, Sally Ralph, Eric Messens, and Wout Boerjan*

Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology, Ghent University, B–9052 Ghent, Belgium (K.M., G.G., E.M., W.B.); United States Dairy Forage Research Center, Agricultural Research Service, United States Department of Agriculture, and Department of Forestry, University of Wisconsin, Madison, Wisconsin 53706 (J.R., H.K., F.L.); and United States Forest Products Laboratory, United States Department of Agriculture Forest Service, Madison, Wisconsin 53705 (S.R.)

Lignin is an aromatic heteropolymer, abundantly present in the walls of secondary thickened cells. Although much research has been devoted to the structure and composition of the polymer to obtain insight into lignin polymerization, the low-molecular weight oligolignol fraction has escaped a detailed characterization. This fraction, in contrast to the rather inaccessible polymer, is a simple and accessible model that reveals details about the coupling of monolignols, an issue that has raised considerable controversy over the past years. We have profiled the methanol-soluble oligolignol fraction of poplar (*Populus* spp.) xylem, a tissue with extensive lignification. Using liquid chromatography-mass spectrometry, chemical synthesis, and nuclear magnetic resonance, we have elucidated the structures of 38 compounds, most of which were dimers, trimers, and tetramers derived from coniferyl alcohol, sinapyl alcohol, their aldehyde analogs, or vanillin. All structures support the recently challenged random chemical coupling hypothesis for lignin polymerization. Importantly, the structures of two oligomers, each containing a γ-*p*-hydroxybenzoylated syringyl unit, strongly suggest that sinapyl *p*-hydroxybenzoate is an authentic precursor for lignin polymerization in poplar.

Lignin is an aromatic heteropolymer that is mainly present in the walls of secondary thickened cells, where it provides strength and impermeability, allowing transport of water and solutes through the vascular system. There is wide interest in understanding the process of lignin biosynthesis and deposition because of its economic relevance; during chemical pulping, lignin needs to be extracted from the wood chips, a process that is expensive and environmentally hazardous. In addition, lignin limits the digestibility of forages. Hence, plant varieties with altered lignin contents may have improved performance as fodder crops or in the production of pulp and paper (Guo et al., 2001; Pilate et al., 2002; Baucher et al., 2003; Boudet et al., 2003).

In dicotyledonous plants, the lignin polymer is made predominantly from the monolignols coniferyl and sinapyl alcohol (Baucher et al., 1998), whereas the lignin of gymnosperms, on the other hand, lacks sinapyl alcohol. After their synthesis, the lignin monomers are transported to the cell wall where they are polymerized in a combinatorial fashion by free-radical coupling mechanisms, generating a variety of structures within the polymer (Boerjan et al., 2003; Ralph et al., 2004b).

By means of a number of chemical degradation methods, such as derivatization followed by reductive cleavage (Lu and Ralph, 1997), acidolysis (Lundquist, 1992), and thioacidolysis (Rolando et al., 1992), and spectroscopic techniques such as NMR (Ralph et al., 1999a; Lu and Ralph, 2003) and Fourier-transform infrared spectroscopy (Faix, 1986), the nature of the chemical bonds and their relative abundance in the final polymer has been elucidated (Adler, 1977; Brunow et al., 1999; Ralph et al., 2004b). However, during lignin polymerization, a fraction of lower M_r phenolic compounds is produced that has escaped a detailed characterization, despite the early use of in vitro dehydropolymerization to obtain low-M_r oligomers for characterization (Freudenberg and Neish, 1968). The study of this plant phenolic fraction is important to better understand lignin polymerization and deposition and to answer some pertinent questions about monolignol coupling in vivo.

The first step in lignin polymerization involves the dehydrogenation of the monolignols by oxidative enzymes, such as peroxidases or laccases, with the formation of radicals (Christensen et al., 2000). Resonance

[1] This work was supported by the Fund for Scientific Research-Flanders (G.0040.00N) and by the European Commission program (QLK5–CT–2000–01493 to E.M. and W.B.), in part by the Department of Energy Biosciences program (DE–AI02–00ER15067) and by the U.S. Department of Agriculture-National Research Initiatives (2001–02176 to J.R.), and by the Instituut voor de aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen (predoctoral fellowship to K.M.).

* Corresponding author; e-mail wout.boerjan@psb.ugent.be; fax 32–9–331–3809.

[w] The on-line version of this article contains Web-only data.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.049304.

stabilizes the radical with unpaired electron density at the C_1, C_3, 4–O, C_5, and C_8 positions (Fig. 1). According to the conventional random coupling hypothesis, the monomeric radicals couple according to their relative supply and coupling propensities, and these reactions are influenced by the macromolecular environment of the cell wall, finally leading to a racemic polymer (Freudenberg and Neish, 1968; Grabber et al., 1996; Ralph et al., 1999b, 2004b; Syrjänen and Brunow, 2000).

A new class of so-called dirigent proteins that are capable of guiding the stereospecific coupling of two coniferyl alcohol radicals into the lignan (+)-pinoresinol has been described (Davin et al., 1997). Lignans are typically optically active compounds, thought to serve as defense substances in plants and derived from the very same monolignols used to generate the lignin polymer (Lewis and Davin, 1999; Sakakibara et al., 2003; Umezawa, 2004). However, the discovery of dirigent proteins has led to a controversial, but widespread, hypothesis that lignin polymerization is tightly controlled by protein-mediated coupling reactions (Lewis, 1999; Chen and Sarkanen, 2004). Although only one dirigent protein, catalyzing the formation of (+)-pinoresinol, has been functionally characterized so far (Davin et al., 1997), the large size of the *DIRIGENT* gene families in a variety of species has been used as an argument that the other linkages in lignin are also protein mediated (Lewis, 1999), although no functional proof has supported this hypothesis yet. A dirigent protein has been immunolocalized to the cambial region and the cell wall (Burlat et al., 2001) and a *DIRIGENT* gene is highly expressed in the lignifying zone of poplar (*Populus* spp.; Hertzberg et al., 2001), corroborating an important role for dirigent proteins in these tissues.

As a first step in deepening our understanding of monolignol coupling and polymerization, and in discriminating lignin from lignan biosynthesis, we reasoned that the structures of low-M_r oligolignols should reflect the in vivo coupling conditions. Hence, we

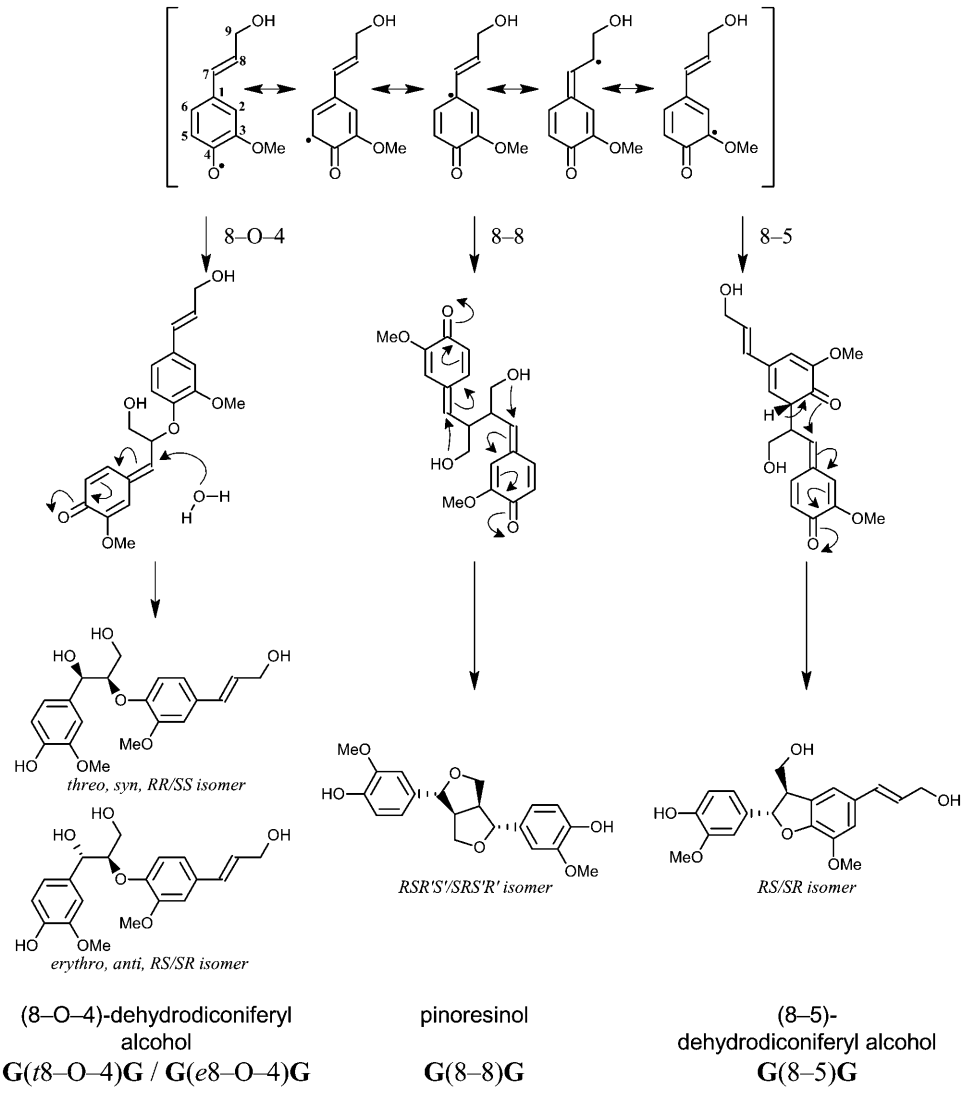


Figure 1. Dilignol formation. Radical-radical coupling involving the C_8 position of a coniferyl alcohol and the 4–O, C_8, or C_5 position of another coniferyl alcohol leading to 8–O–4, 8–8, and 8–5 linkages. The quinone methide intermediates are subsequently rearomatized to (8–O–4)-dehydrodiconiferyl alcohol, pinoresinol, and (8–5)-dehydrodiconiferyl alcohol. The R/S nomenclature specifies carbons 7 and 8, namely RS = 7R,8S.

characterized this fraction in poplar xylem, a tissue that is heavily lignified. In transgenic poplars with reduced lignin content, this oligolignol fraction was severely depleted. We identified the structures of 38 phenolic compounds, most of which were dimeric, trimeric, or tetrameric oligolignols derived from coniferyl and sinapyl alcohols and their aldehydes. In addition, the structures of two compounds demonstrate that sinapyl *p*-hydroxybenzoate has to be considered as an authentic lignin precursor in poplar. The structures of all identified compounds are in accordance with the recently challenged combinatorial coupling hypothesis. This is the first study to our knowledge describing the low-M_r oligolignol fraction from lignifying tissue.

RESULTS

Characterization of Oligolignols from Lignifying Poplar Xylem

Our aim was to obtain insight into the process of monolignol coupling in the cell wall by characterizing the chemical structures of a large number of low-M_r, monolignol-coupling products, and to investigate whether these structures are consistent with a combinatorial coupling process under chemical control. Because monolignol coupling occurs excessively during lignin polymerization, such a low-M_r oligolignol fraction is expected to be present in the walls of lignifying cells. To identify this oligolignol fraction, we profiled the methanol-soluble phenolics present in xylem extracts of wild-type and caffeoyl-CoA *O*-methyltransferase (CCoAOMT)-deficient poplars by HPLC (Fig. 2). Because xylem of the latter plants accumulates less lignin (Meyermans et al., 2000; Zhong et al., 2000), it is an ideal material to identify this fraction, because the oligolignols are expected to be less abundant. Indeed, in the last half of the chromatogram between 11 and 24 min (Fig. 2), a family of compounds abundantly present in wild-type poplar was barely detectable in HPLC profiles of poplars down-regulated for CCoAOMT, suggesting that their synthesis involved the monolignol biosynthesis pathway. A similar HPLC profile, showing a depletion of peaks in the second half of the chromatogram was also observed for transgenic poplars down-regulated for cinnamoyl-CoA reductase (J.-C. Leplé, K. Morreel, C. Lapierre, K. Ruel, J.-P. Joseleau, G. Goeminne, R. De Rycke, E. Messens, G. Pilate, and W. Boerjan, unpublished data). All these compounds had similar UV/visible (Vis) adsorption spectra with a maximum at approximately 270 nm. Interestingly, although the large fraction of peaks in the last half of the chromatogram was almost absent in the xylem of CCoAOMT-deficient plants, the total peak height (the sum of the heights of all peaks in a chromatogram, divided by the dry weight of xylem tissue) was 2.4-fold higher than that of the wild type. This increase could be attributed primarily to three newly accumulating compounds (the phenolic glucosides of vanillin, caffeic acid, and sinapic acid) that had been identified previously (Meyermans et al., 2000). These three glucosides are thought to be detoxification and/or storage products of the free acids that accumulate as a consequence of a redirection of the flux of caffeoyl-CoA to caffeic acid and further to sinapic acid, rather than to feruloyl-CoA (Meyermans et al., 2000).

To identify the structures of the presumed oligolignols in wild-type poplar, HPLC fractions (Fig. 3, chromatogram D and table) of the complete set of peaks in the second half of the chromatogram were collected and separated on liquid chromatography-mass spectrometry (LC-MS) for further structural elucidation. By mass spectrometry/mass spectrometry (MS/MS) analysis, a tentative structure was proposed. The assigned structure for a number of peaks could be authenticated by spiking and MS/MS analysis of the synthesized compound. Several compounds were trivially assigned by analogy with confirmed peaks, solely from their mass spectral data. As presented in Figure 4, 38 oligolignols were authen-

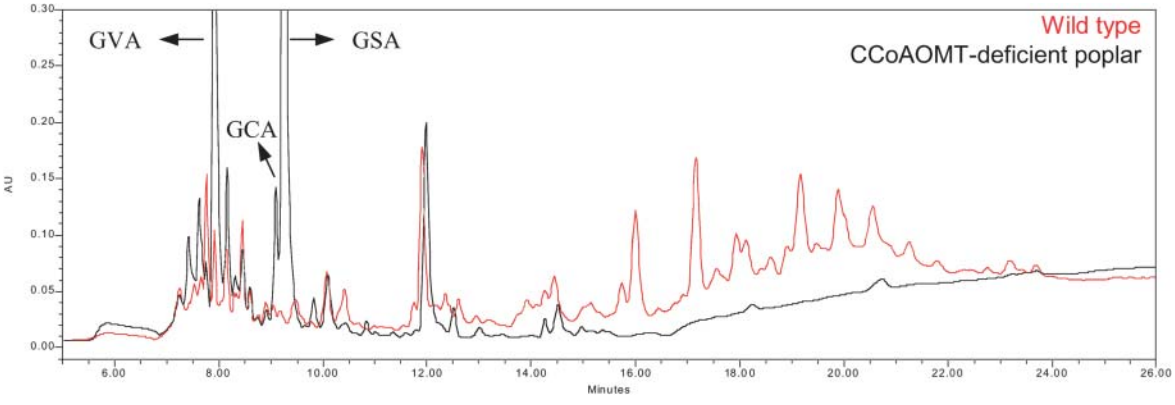
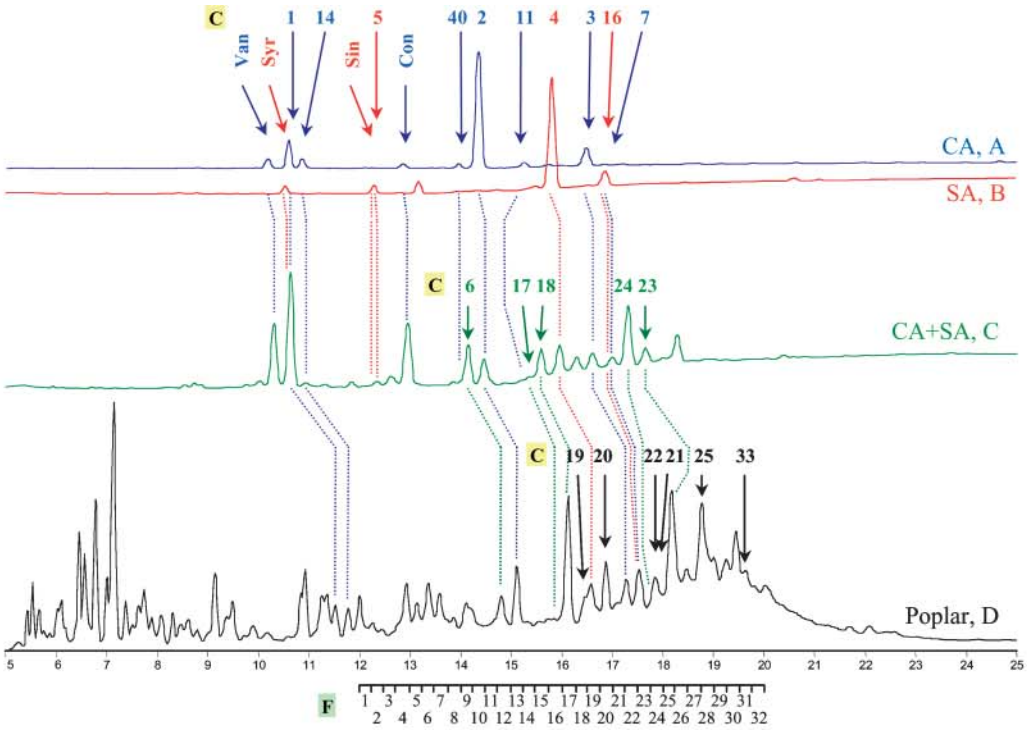


Figure 2. HPLC chromatogram overlay. Chromatograms from xylem extracts of a wild-type poplar (red) and a transgenic CCoAOMT-deficient poplar (black) are shown. GVA, O^4-β-D-glucopyranosyl-vanillic acid; GCA, O^3-β-D-glucopyranosyl-caffeic acid; GSA, O^4-β-D-glucopyranosyl-sinapic acid.



C	M	t_R	F	Conc m	Conc sem	Shorthand name	Trivial name
1	v	11.5		31.3	12	G(*t*8–O–4)G	*threo*-(8–O–4)-dehydrodiconiferyl alcohol
2	v	15.1		190	70	G(8–5)G	(8–5)-dehydrodiconiferyl alcohol
3	v	17.3		70.2	26	G(8–8)G	pinoresinol
4	v	16.6		29.1	7	S(8–8)S	syringaresinol
5	v		6			S(*t*8–O–4)S	*threo*-(8–O–4)-dehydrodisinapyl alcohol
6	v	14.8		57.3	18	S(8–5)G	simulanol
7	v	17.5		13.0	4	G(8–5)G′	balanophonin
8			5			G(*t*8–O–4)S	*threo*-guaiacylglycerol 8–O–4 sinapyl ether
9			12			S(8–O–4)S′	
10			15			G(8–O–4)S′ / S(8–O–4)G′	
11	v		15			G(8–5)V′	ficusal
12			21			S(8–5)G′	
13			22			G(8–O–4)G′	
14	v	11.8		21.2	9	G(*t*8–O–4)G(*t*8–O–4)G	
15			18			G(*t*8–O–4)G(8–8)G	
16	v	17.4		n.d.	n.d.	S(*t*8–O–4)S(8–8)S	*threo*-buddlenol D
17	v	15.8		10.5	3	S(*t*8–O–4)S(8–5)G	
18	v	16.1		168	40	G(*t*8–O–4)S(8–5)G	*threo*-buddlenol B
19		16.4		21.0	13	SP(8–8)S	
20	v	16.9		55.3	16	G(*e*8–O–4)S(8–5)G	*erythro*-buddlenol B
21		17.8		31.2	3	S(*t*8–O–4)S(8–5)G′	
22		18.0		n.d.	n.d.	G(*t*8–O–4)S(8–8)G	*threo*-buddlenol E
23	v	18.2		123	15	G(*t*8–O–4)S(8–5)G′	*threo*-buddlenol A
24	v	18.3		n.d.	n.d.	G(*t*8–O–4)S(8–8)S	*threo*-buddlenol C
25	v	18.8		85.5	10	G(*e*8–O–4)S(8–5)G′	*erythro*-buddlenol A
26			16			G(8–5)G(*t*8–O–4)G′	
27			17			G(*t*8–O–4)G(8–5)V′	
28			20			G(*t*8–O–4)G(8–5)G′	
29			21			G(*e*8–O–4)G(8–5)G′	
30			22			S(8–O–4*)S(8–5)G	
31			27			G(*t*8–O–4)S′(8–8)S / G(*t*8–O–4)S(8–8)S′	
32			27			G(*e*8–O–4)S(8–8)S	*erythro*-buddlenol C
33	v	19.6		62.2	4	G(8–O–4)S(8–8)S(8–O–4)G	hedyotisol
34			22			S(*t*8–O–4)SP(8–8)S / S(*t*8–O–4)S(8–8)SP	
35			27			G(*e*8–O–4)G(8–O–4)S(8–8)S′ / G(*e*8–O–4)G(8–O–4)S′(8–8)S	
36			22			G(8–O–4)G(8–8)S(8–O–4)G	
37			25			S(8–O–4)S(8–8)S(8–O–4)G	
38			23			G(*t*8–O–4)G(8–O–4)S(8–8)S	
39						S(*t*8–O–4)G(8–5)G	
40	v					G(*t*8–O–4)G(8–5)G	

Figure 3. (*Legend appears on following page.*)

ticated or tentatively identified in this way (see supplemental data, available at www.plantphysiol.org, for the MS/MS spectra of all identified compounds and the arguments for the assignment of a particular structure).

The oligomers were mainly composed of units derived from coniferyl alcohol (guaiacyl, **G**) from sinapyl alcohol (syringyl, **S**) and from coniferaldehyde (**G′**), and a few contained units derived from sinapaldehyde (**S′**), vanillin (**V′**), and sinapyl *p*-hydroxybenzoate (**SP**; for nomenclature, see ''Materials and Methods''). All possible 8–O–4-, 8–5-, and 8–8-coupled homodimers of **G** and **S** units (Fig. 4), i.e. **G**(*t*8–O–4)**G**, **1**; **G**(8–5)**G**, **2**; **G**(8–8)**G**, **3**; **S**(8–8)**S**, **4**; and **S**(*t*8–O–4)**S**, **5**, were detected. In addition to **G** and **S** units, alternative units, such as **G′**, **S′**, **V′**, and **SP** were found in the heterodimeric fraction, i.e. **S**(8–5)**G**, **6**; **G**(8–5)**G′**, **7**; **G**(*t*8–O–4)**S**, **8**; **S**(8–O–4)**S′**, **9**; **G**(8–O–4)**S′** or **S**(8–O–4)**G′** (compound **10**); **G**(8–5)**V′**, **11**; **S**(8–5)**G′**, **12**; **G**(8–O–4)**G′**, **13**; and **SP**(8–8)**S**, **19**. Homodimers of **G′**, **S′**, **V′**, and **SP** were not detected, presumably because of the low abundance of their precursor monomers. Except for **G**(8–5)**G**(*t*8–O–4)**G′**, **26**, all tri- and tetrameric compounds were composed of a **G** or **S** unit linked by a β-aryl ether bond to a moiety derived from one of the dimers mentioned above (Figs. 3 and 4), or to an **S**(8–8)**G** or **S**(8–8)**S′** moiety, i.e. **G**(*t*8–O–4)**G**(*t*8–O–4)**G**, **14**; **G**(*t*8–O–4)**G**(8–8)**G**, **15**; **S**(*t*8–O–4)**S**(8–8)**S**, **16**; **S**(*t*8–O–4)**S**(8–5)**G**, **17**; **G**(*t*8–O–4)**S**(8–5)**G**, **18**; **G**(*e*8–O–4)**S**(8–5)**G**, **20**; **S**(*t*8–O–4)**S**(8–5)**G′**, **21**; **G**(*t*8–O–4)**S**(8–8)**G**, **22**; **G**(*t*8–O–4)**S**(8–5)**G′**, **23**; **G**(*t*8–O–4)**S**(8–8)**S**, **24**; **G**(*e*8–O–4)**S**(8–5)**G′**, **25**; **G**(*t*8–O–4)**G**(8–5)**V′**, **27**; **G**(*t*8–O–4)**G**(8–5)**G′**, **28**; **G**(*e*8–O–4)**G**(8–5)**G′**, **29**; **G**(*t*8–O–4)**S**(8–8)**S′** or **G**(*t*8–O–4)**S′**(8–8)**S** (compound **31**); **G**(*e*8–O–4)**S**(8–8)**S**, **32**; **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33**; **S**(*t*8–O–4)**SP**(8–8)**S** or **S**(*t*8–O–4)**S**(8–8)**SP** (compound **34**); **G**(*e*8–O–4)**G**(8–O–4)**S′**(8–8)**S** or **G**(*e*8–O–4)**G**(8–O–4)**S**(8–8)**S′** (compound **35**); **G**(8–O–4)**G**(8–8)**S**(8–O–4)**G**, **36**; **S**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **37;** and **G**(*t*8–O–4)**G**(8–O–4)**S**(8–8)**S** **38**. Compound **30**, **S**(8–O–4*)**S**(8–5)**G**, is likely formed by simple benzylic oxidation of the trimer **S**(8–O–4)**S**(8–5)**G**, **17**.

All detected tetramers were derived from an 8–8-dimeric moiety to which **G** and/or **S** units were attached. The biosynthesis of these compounds is initiated by monomer-monomer coupling, yielding an 8–8-linked dimer with two phenolic groups that are amenable to oxidation by peroxidase/H_2O_2, for instance. Hence, further chain extension might be initiated at either phenol of this dimer, yielding tetramers characterized by an internal 8–8-linked unit or an 8–8-linked end group. Higher order oligomers (pentamers, hexamers, etc.) might be present in poplar xylem as well; some of the corresponding m/z values were found by LC-MS analysis, but they were present in minute amounts.

If the production of these oligolignols solely depended on the chemical coupling conditions in the cell wall, their concentrations would be in accordance with the relative supply and cross-coupling propensities of the monomers. Therefore, the concentrations were estimated for the identified oligolignols based on the HPLC chromatograms of the xylem extracts (Fig. 3). Fifteen of the identified compounds were separated sufficiently and abundantly allowing their pseudo-quantification. Together, these 15 oligolignols accounted for approximately 0.05% of the dry weight of xylem tissue. The major detected dilignol was (8–5)-dehydrodiconiferyl alcohol, **G**(8–5)**G**, **2**, whereas the major trilignols were *threo*-buddlenol B, **G**(*t*8–O–4)**S**(8–5)**G**, **18**, and its corresponding cinnamaldehyde, i.e. *threo*-buddlenol A, **G**(*t*8–O–4)**S**(8–5)**G′**, **23**. The *erythro*-isomers of these trilignols accounted for 25% and 42% of the total amounts (*threo* + *erythro*) of buddlenol B and A, respectively. The only tetralignol that could be quantified was **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33**. Overall, taking the concentrations into account, the quantified oligolignols were composed mainly of **G** (59%), **S** (31%), and **G′** (10%) units and traces of **V′**, **S′**, and **SP** units and were linked by 8–5 (47%), 8–O–4 (42%), and 8–8 (11%) bonds. No **H** units were detected in any of the coupling products.

Oligolignol Profiling of Synthetic Mixtures

Our hypothesis is that the oligolignols are derived from phenolic units through oxidation, followed by chemical coupling that is not protein mediated. Thus, synthetic reaction mixtures, prepared by the oxidation of coniferyl alcohol, sinapyl alcohol, or both coniferyl and sinapyl alcohols, resulting in **G**, **S**, or **G** + **S** synthetic oligolignol mixtures, respectively, should reveal the same oligolignol structures as those detected in poplar xylem extracts. These oligolignol mixtures were prepared and separated with the same reversed-phase HPLC method and compared to the oligolignol profiles obtained from the poplar xylem extracts (Fig. 3).

Figure 3. Oligolignol profile. HPLC chromatograms of oligomers present in mixtures obtained from the oxidation of coniferyl alcohol (CA; A), sinapyl alcohol (SA; B), and coniferyl and sinapyl alcohols (CA + SA; C), and of a methanol xylem extract from poplar (poplar; D). Oligomers identified in chromatograms A to D are shown in blue, red, green, and black, respectively. Vertical dashed lines indicate the corresponding positions of the identified peaks in the different chromatograms. Isolated HPLC fractions (F) for LC-MS/MS analysis are shown below the retention time. Compound number (C), whether present or not in the **G** + **S** synthetic mixture (M), retention time (t_R), or HPLC fraction where the compound was detected, concentrations (Conc; pmol mg^{-1} dry weight of xylem tissue), shorthand, and trivial names of the oligolignols are given below. Retention times are based on the chromatogram of the xylem extract (D). m, mean; n.d., not determined due to coelution; sem, SE of the mean. The presence of **G**(*e*8–O–4)**S**(8–5)**G**, **20**; **G**(*e*8–O–4)**S**(8–5)**G′**, **25**; and **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33** in the synthetic mixtures was only established by NMR analysis following their purification.

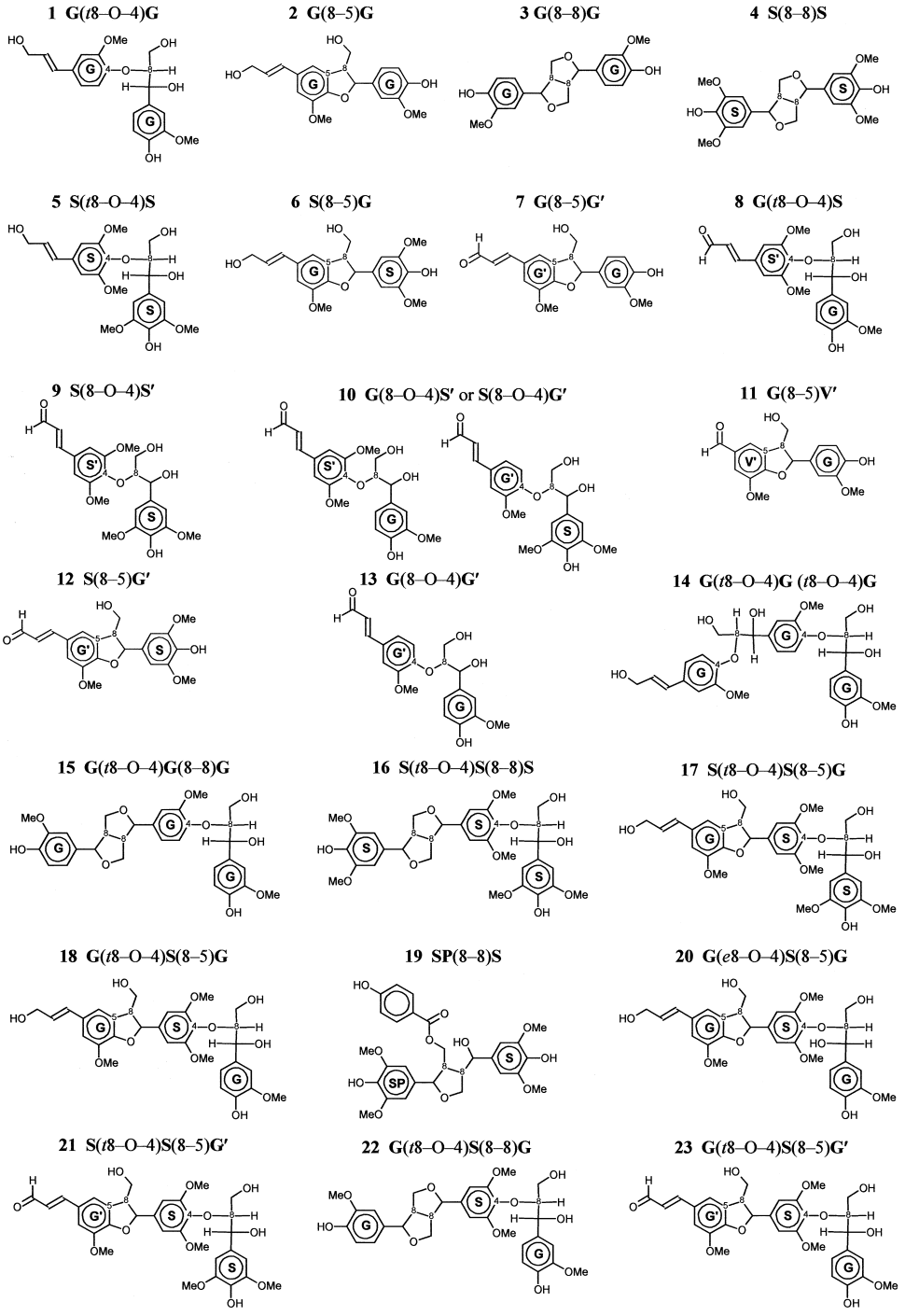


Figure 4. Oligolignol structures present in poplar xylem extracts. Molecular structure and shorthand naming inferred from the spectra are shown. (*Figure continues on following page.*)

The chromatogram of the synthetic **G** oligolignol mixture showed the main types of dimerization products involving the C_8 position, i.e. **G**(*t*8–O–4)**G**, **1**, **G**(8–5)**G**, **2**, and **G**(8–8)**G**, **3**. In accordance with the in vivo situation, two phenylcoumaran dimers were detected for which a **G** unit was connected to a unit derived from coniferaldehyde or vanillin, i.e. **G**(8–5)**G**′, **7** and **G**(8–5)**V**′, **11**. Although the coniferyl alcohol used for the synthetic mixture was virtually pure, both free coniferaldehyde and vanillin were present as well in the **G** oligolignol mixture based on their MS/MS spectra and the spiking of synthetic products. This indicates that coniferyl alcohol is oxidized to aldehydes under the synthetic conditions. Two trimers were found, namely **G**(*t*8–O–4)**G**(*t*8–O–4)**G**, **14** and **G**(*t*8–O–4)**G**(8–5)**G** (compound **40**; Fig. 3), the latter of which was not detected in the xylem extracts. No higher order oligomers were detected. Oligolignol units were mainly 8–5 linked (Fig. 3).

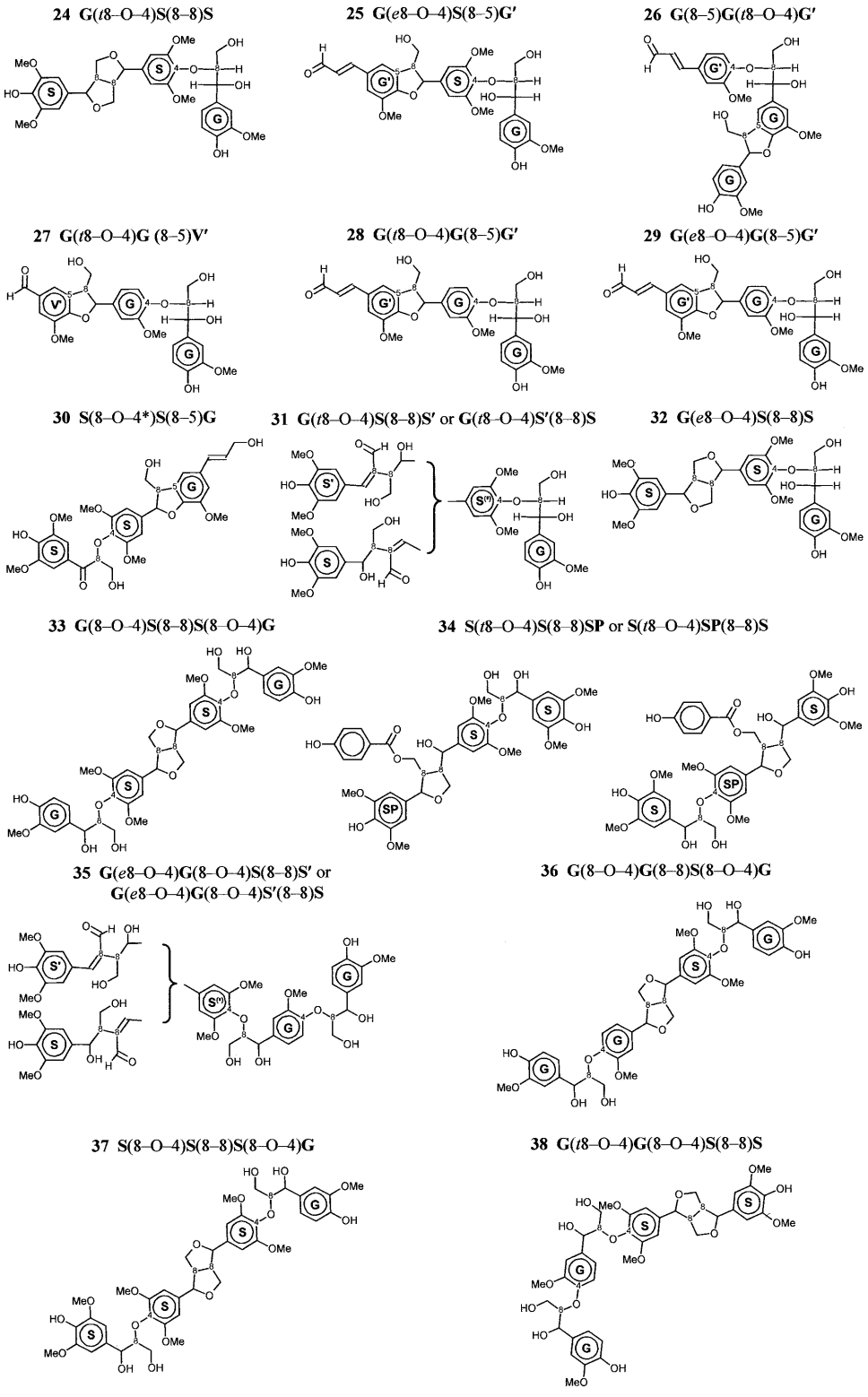


Figure 4. (*Continued.*)

HPLC analysis of the synthetic **S** oligolignol mixture showed the presence of both **S**(8–8)**S**, **4** and **S**(*t*8–O–4)**S**, **5** dimers, and only one trimer, **S**(*t*8–O–4)**S**(8–8)**S**, **16**. **S**(8–8)**S**, **4** was the major compound in this synthetic oligolignol mixture (Fig. 3). By MS/MS analysis and the spiking of standards, two peaks were identified as free sinapaldehyde and syringaldehyde, although the oligolignol mixture was prepared starting from virtually pure sinapyl alcohol.

All compounds identified in the **G** or **S** oligolignol mixtures were also found in the synthetic **G** + **S** oligolignol mixture, which, in addition, contained the **S**(8–5)**G**, **6** dimer, its β-aryl ether-derived trimers, i.e. **S**(*t*8–O–4)**S**(8–5)**G**, **17** and **G**(*t*8–O–4)**S**(8–5)**G**, **18**, the corresponding aldehyde of the latter trimer, **G**(*t*8–O–4)**S**(8–5)**G′**, **23**, and the trimer where a **G** unit is connected via a β-aryl ether to a syringaresinol substructure, **G**(*t*8–O−4)**S**(8–8)**S**, **24**. The latter was the most abundant oligolignol present in the **G** + **S** synthetic mix (Fig. 3).

All identified peaks in the synthetic mixtures were detected in poplar xylem extracts, except for the poorly abundant **G**(*t*8–O–4)**G**(8–5)**G**, **40**, vanillin, and syringaldehyde. Compared to the synthetic mixtures, xylem extracts contained some additional oligolignols, especially tri- and tetralignols (Figs. 3 and 4). The *erythro*-diastereomers of the more abundant xylem oligolignols, such as **G**(*e*8–O–4)**S**(8–5)**G**, **20** and **G**(*e*8–O–4)**S**(8–5)**G′**, **25**, and the tetralignol **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33**, were clearly detected in the HPLC profiles of xylem extracts, but their presence in the synthetic mixtures was only established by NMR analysis following their purification as *threo*/*erythro* mixtures; *threo*-β-aryl ethers are strongly favored in the borate buffer system used (Landucci et al., 1995).

DISCUSSION

The Oligolignol Structures of Poplar Xylem Extracts Are in Agreement with Chemical Coupling Reactions

We have characterized the methanol-soluble oligolignol fraction of poplar xylem to investigate whether their structures are consistent with a chemical coupling process. UV/Vis and MS/MS spectra were used for the initial elucidation of the structure of these compounds. For many of them, the proposed structures were subsequently validated by spiking and MS/MS analysis of synthesized reference compounds. We have authenticated or tentatively identified the structures of 38 compounds, most of which correspond to simple coupling products of monolignols, including dimers, trimers, and tetramers. All structures suggest they correspond with products of radical coupling reactions, and no further modifications invoking enzymatic reactions were evidenced.

The high frequency of β-aryl ether units in trimers and tetramers (all composed of units 8–O–4 linked to an 8–O–4-, 8–5-, or 8–8-coupled dimer) is in agreement with the chemical cross-coupling reactions between a monomer and an oligomer. There are only two possibilities for a hydroxycinnamyl alcohol to couple at its favored C_8 position with a **G** phenolic end group (at its 4–O or C_5 position) and only one for coupling with an **S** phenolic end group (at its 4–O position) (Boerjan et al., 2003; Ralph et al., 2004b).

A survey of the β-aryl ether units in the trimers and tetramers shows that 8–O–4 coupling occurs between coniferyl alcohol and both **G** and **S** units, whereas sinapyl alcohol forms only 8–O–4 linkages to **S** units; this observation is again in agreement with a nonprotein-mediated chemical coupling reaction, where the 8–O–4 radical coupling propensities do not favor a reaction between sinapyl alcohol and a **G** unit because of factors, such as oxidation potential and radical reactivity (Landucci et al., 1992; Syrjänen and Brunow, 1998). To further support this hypothesis, the compound **S**(*t*8–O–4)**G**(8–5)**G**, **39**, which is the **S**-type β-aryl ether of the most abundant dimer in the xylem extract, **G**(8–5)**G**, **2**, was synthesized and searched for by both HPLC analysis and LC-MS analysis of isolated HPLC fractions. This compound, **S**(*t*8–O–4)**G**(8–5)**G**, **39** was found neither in poplar xylem extracts nor in synthetic oligolignol mixtures (see below).

The Structures of the Oligolignols Are in Agreement with Endwise Polymerization Conditions in the Cell Wall

As is the case for lignification, the oligolignols described here are likely produced by an endwise rather than by a bulk polymerization process. Zulauf-verfahren dehydrogenation polymer (DHP) reactions, which mimic a bulk polymerization process (Freudenberg, 1959), have shown that monolignol radicals prefer to couple with like monolignol radicals rather than to form heterodimers or to cross-couple with dimers or higher oligomers. Hence, the detection of both heterodimers and heterooligomers in the xylem oligolignol fraction supports an endwise polymerization process. Furthermore, bulk polymerization results in oligomer-oligomer couplings, producing 5–5- and 4–O–5-linked structures (Sarkanen, 1971), which are not detected in the xylem oligolignol mixture. On the other hand, endwise polymerization, mimicked by Zutropf-verfahren DHP reactions, results from the gradual supply of monomers to the site of polymerization and represses especially the 8–8-coupling mode (Grabber et al., 1996; Syrjänen and Brunow, 2000). In the xylem oligolignol fraction, only 11% of the linkages were 8–8, a frequency that is in agreement with Zutropf-verfahren DHP reactions. Together, the oligolignol structures in the xylem extract are in agreement with coupling conditions favoring endwise coupling.

The β-Aryl Ether Units in Oligolignols Are Mainly *threo*-Diastereomers

The characterization of the oligolignols present in poplar xylem showed that 14 of the 25 dimeric, trimeric, and terminal tetrameric β-aryl ethers were present only in the *threo* configuration, whereas the β-aryl end group of compound **35**, **G**(*e*8–O–4)**G**(8–O–4)**S**(8–8)**S′** or **G**(*e*8–O–4)**G**(8–O–4)**S′**(8–8)**S**, was present in the *erythro* form. The *threo*/*erythro* configuration could not be determined from the MS/MS spectra of **S**(8–O–4)**S′**, **9**; **G**(8–O–4)**S′** or **S**(8–O–4)**G′**, **10**; and **G**(8–O–4)**G′**, **13**, and from the MS/MS spectra of the tetralignols with an internal 8–8-linked moiety, i.e. **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33**; **G**(8–O–4)**G**(8–8)**S**(8–

O–4)**G**, **36**; and **S**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **37**. Both *threo*- and *erythro*-diastereomers were detected for the remaining 4 of the 25 β-aryl ethers, i.e. for the structures **G**(*t*8–O–4)**S**(8–8)**S**, **24** and **G**(*e*8–O–4)**S**(8–8)**S**, **32**, and **G**(*t*8–O–4)**G**(8–5)**G′**, **28** and **G**(*e*8–O–4)**G**(8–5)**G′**, **29**, and the more abundant trilignols, i.e. **G**(*t*8–O–4)**S**(8–5)**G**, **18** and **G**(*e*8–O–4)**S**(8−5)**G**, **20**, **G**(*t*8–O–4)**S**(8–5)**G′**, **23** and **G**(*e*8–O–4)**S**(8–5)**G′**, **25**. These four compounds could be clearly quantified in the HPLC profile of xylem extracts; the *threo* forms were present for 58% and 75% of the total amount of **G**(*t*8–O–4)**S**(8–5)**G**, **18** and **G**(*e*8–O–4)**S**(8–5)**G**, **20**; and **G**(*t*8–O–4)**S**(8–5)**G′**, **23** and **G**(*e*8–O–4)**S**(8–5)**G′**, **25** with *threo*/*erythro* ratios of 3:2 and 3:1, respectively. Taken together, the *threo*-diastereomers are clearly the most present among the β-aryl ethers.

Both *threo*- and *erythro*-β-aryl ethers are also found in lignin, but, in contrast to the xylem-extracted oligolignols, the *erythro* forms of the β-aryl ether linkages predominate in angiosperm lignins (Brunow et al., 1993; Akiyama et al., 2003). Because gymnosperm lignins, composed of **G** units, contain approximately equal amounts of *threo* and *erythro* configurations and because the *threo*/*erythro* ratios correlate inversely with the **S**:**G** ratios in dicots (Akiyama et al., 2003), the preponderance of the *erythro* form has been attributed to the presence of β-syringyl ether structures in angiosperm lignins. Both in vivo and in vitro, 8–O–4-guaiacyl ethers and 8–O–4-syringyl ethers in lignin are produced with approximately 50:50 and approximately 25:75 *threo*/*erythro* ratios, respectively, whereas their equilibrium ratios are nearly equal (Brunow et al., 1993; Ralph et al., 2004b). The reason for the apparent *threo*-isomer predominance in the xylem oligolignol fraction is currently not clear, unless for some reason *erythro*-isomers couple (to higher oligomers) more rapidly.

Radical Coupling Reactions Accept Alternative Monomers

Besides traditional **G** and **S** units, some oligolignols contain alternative units, such as **G′**, **S′**, **V′**, and **SP**. Importantly, the structures of a few of these oligolignols, namely the trimers, imply that these alternative units arise from the coupling of the corresponding monomers rather than from postcoupling oxidation or derivatization reactions.

For example, cross-coupling of sinapaldehyde appears to result in **S**(8–8)**S′**, which, after further coupling with sinapyl alcohol, results in compound **31**, i.e. **G**(*t*8–O–4)**S**(8–8)**S′** or **G**(*t*8–O–4)**S′**(8–8)**S** and compound **35**, i.e. **G**(*e*8–O–4)**G**(8–O–4)**S′**(8–8)**S** or **G**(*e*8–O–4)**G**(8–O–4)**S**(8−8)**S′** (Fig. 4). In constrast to the 8–8 coupling of two sinapyl alcohol radicals to **S**(8–8)**S 4**, with two tetrahydrofuran rings, no ring structures are formed during the 8–8 coupling of sinapyl alcohol with a cinnamaldehyde. The 8–8 coupling between two sinapyl alcohol radicals forms a bis-quinone methide intermediate. Each quinone methide is rearomatized by internal nucleophilic attack of the 9–OH of the other unit resulting in a resinol unit (Fig. 1). However, when one of the C_9 positions in the dimer is oxidized or derivatized, it is no longer available to trap the quinone methide of the other unit. In this case, rearomatization of the other unit can only proceed by the nucleophilic attack of, for example, an incoming water molecule and no tetrahydrofuran ring is formed (Lu and Ralph, 2002). The quinone methide derived from the sinapaldehyde unit is rearomatized by the elimination of the C_8 proton, regenerating the enone function, rather than by a nucleophilic attack at C_7. It should be noted that attempted cross-coupling of coniferaldehyde or sinapaldehyde with normal monolignols did not result in any (**G′** or **S′**)(8–O–4)(**G** or **S**) products; the only cross-product isolated was **S**(8–O–4)**S′** (H. Kim, unpublished data). Therefore, structures **31** and **35** remain unauthenticated, but, if correct, indicate that sinapaldehyde is involved directly in the coupling reactions, i.e. that sinapaldehyde is the monomer for this moiety, and not sinapyl alcohol.

The cinnamaldehyde monomers themselves can either be the reaction products of cinnamoyl-CoA reductase, which are transported to the cell wall as aldehydes, or be derived from the precoupling oxidation of the cinnamyl alcohols already present in the cell wall. Because the cinnamaldehyde-derived units are higher in lignins of transgenic plants down-regulated for cinnamyl alcohol dehydrogenase than those in lignins of wild-type plants (Baucher et al., 1996; Kim et al., 2003), these units in the xylem oligolignol fraction are probably made from coniferaldehyde that is synthesized within the cell and transported to the cell wall. However, we cannot exclude that at least part of the **G′**, **S′**, and **V′** units are derived from the oxidation of the monolignols in the cell wall prior to cross-coupling. For example, although difficult to extrapolate to the in vivo situation, coniferaldehyde, sinapaldehyde, and vanillin were also found in the synthetic mixtures made from coniferyl and sinapyl alcohols, in addition to the coupling products; peroxidase/H_2O_2 also causes such oxidation.

Sinapyl *p*-Hydroxybenzoates Are Precursors in Lignin Biosynthesis

More compelling evidence for the incorporation of alternative units is obtained by the identification of compounds **19** and **34**, i.e. **SP**(8–8)**S** and its sinapyl alcohol coupling product **S**(*t*8–O–4)**SP**(8–8)**S** or **S**(*t*8–O–4)**S**(8–8)**SP**. These results strongly indicate that *p*-hydroxybenzoic acid is esterified by sinapyl alcohol prior to radical cross-coupling, because the product is clearly derived from cross-coupling of sinapyl *p*-hydroxybenzoate with sinapyl alcohol (Lu et al., 2004). Alternatively, **SP**(8–8)**S** is formed by ring opening of **S**(8–8)**S**, followed by the acylation with *p*-hydroxybenzoic acid. This process would require two enzymatic activities, which is less likely, given that the

radical-radical coupling reactions for lignification occur in the cell wall.

γ-Acylated **G** and **S** units have been detected in the lignins of many species. For example, sinapyl acetate is implicated similarly as a monomer in lignification in kenaf bast fibers (Lu and Ralph, 2002). The lignin of grasses is adorned with γ-*p*-coumarate substituents on a variety of lignin units. Lignin structural analyses of poplar, aspen, willow, and palm have shown that γ-*p*-hydroxybenzoate esters are present (Smith, 1955; Nakano et al., 1961; Sun et al., 1999; Meyermans et al., 2000; Landucci and Ralph, 2001). The fact that only **S** units are γ-*p*-hydroxybenzoylated in both poplar lignin and the oligolignol fraction supports the contention that sinapyl *p*-hydroxybenzoate is produced enzymatically and used as an authentic monomer for lignification in poplar.

Are These Oligomers Destined for Lignin?

Taken together, the dilignols, trilignols, and tetralignols described here are produced by radical endwise condensation reactions and no postcoupling enzymatic reactions seem to be involved because no products were detected resulting from further metabolism of the oligolignols. A pathway in which the oligolignols are used as the main building blocks of lignin is considered to be, at best, a minor one because lignins contain relatively few cinnamyl alcohol end groups, indicating that lignin is mainly produced by the addition of monolignols to the growing polymer and not by the concatenation of preformed oligomers. **G**- and **S**-type β-aryl ether and phenylcoumaran dimers can only add to another monolignol or oligolignol via the free-phenolic function; the unsaturated propenol side chain is blocked from any further reactions because peroxidase oxidizes specifically the phenol function and all couplings are radical-radical reactions. Thus, coupling involving substantial amounts of oligolignols would result in a high proportion of terminal alcohol residues in lignin, which is not observed. For example, it was estimated that over 95% of the lignin units are not derived from dimerization reactions, at least in softwood (Hatfield and Vermerris, 2001).

Of the 38 compounds characterized in this study, 20 have been previously identified from a variety of plant species and tissues, but because of their sporadic identification from various species and tissues, most of these oligolignols have been considered as lignans, compounds with a defensive role in plants. It is the large number of oligolignols identified in this study from one species and from a single tissue with extensive lignification and the nature of their chemical structures that allows us to conclude that these compounds have to be considered as a class of monolignol-coupling products that are formed under the ambient monolignol concentrations and oxidative conditions in the cell wall.

Our data support the recently challenged combinatorial chemical coupling hypothesis of monolignols (Ralph et al., 2004b), but do not exclude that the coupling of certain dilignols may be assisted by dirigent proteins. To investigate this possibility, the oligolignols present in xylem extracts should be purified and analyzed by chiral HPLC to determine their enantiomeric ratios. This will give insight into the elusive role of dirigent proteins in oligolignol synthesis. In another article (Morreel et al., 2004), we show that the relative abundance of these oligolignols is altered dramatically in transgenic poplar down-regulated for caffeic acid *O*-methyltransferase and that novel oligolignols, derived from products of incomplete monolignol biosynthesis, are produced.

MATERIALS AND METHODS

Growth Conditions and Plant Material

Wild-type and CCoAOMT down-regulated poplars (*Populus tremula* × *P. alba* clone INRA no. 717–1B4; Meyermans et al., 2000) were propagated in vitro on Murashige and Skoog medium (Duchefa, Haarlem, The Netherlands). Rooted plantlets were transferred to the greenhouse (21°C, 60% humidity, 16-h light/8-h dark regime, 40–60 μmol m^{-2} s^{-1} photosynthetic photon flux) and grown for 3 months until harvest, reaching a height of approximately 1.5 m.

Approximately 300 mg of xylem tissue were harvested from a 10-cm-long, debarked stem (by scraping with a scalpel), cut at 15 cm above ground. After grinding in liquid nitrogen, the tissue was extracted with 15 mL of methanol. The supernatant was subsequently removed and the residue lyophilized and weighed (approximately 70 mg).

HPLC Analysis

An aliquot (1.0 mL) of the methanol phase was lyophilized and extracted with cyclohexane/water acidified with 0.1% trifluoroacetic acid (1:1; v/v), and separated on HPLC with a Luna C18(2) column (250 × 4.6 mm, 5 μm; Phenomenex, Torrance, CA), as previously described (Meyermans et al., 2000). A valley-to-valley integration of the chromatogram was applied using the following parameter values: peak width, 15 s and threshold, 17 μV s^{-1}. Using the maximum absorbance value between 230 and 450 nm, quantification was based on the peak height instead of the peak area as the latter method is more sensitive to impurities (Snyder et al., 1997) and standardized to the dry weight. In addition, the HPLC procedure was carried out on sufficient plant material to collect 0.3-mL fractions that were freeze-dried and redissolved in 0.1 mL 1% aqueous triethylammonium acetate for LC-MS/MS.

For NMR analysis, repetitive HPLC separations were used to collect at least 0.1 mg of the compound of interest, followed by a final repurification on the Luna C18(2) column described above.

LC-MS/MS Analysis

HPLC fractions were injected by means of a SpectraSystem AS1000 autosampler (Thermo Separation Products, Riviera Beach, FL) onto a reversed-phase Luna C18(2) column (150 × 2.1 mm, 3 μm; Phenomenex). A gradient separation (SpectraSystem P1000XR HPLC pump; Thermo Separation Products) was run from 1% aqueous triethylammonium acetate (solvent A, pH 5) to methanol-acetonitrile (25:75; v/v; 1% triethylammonium acetate; solvent B) using the following conditions: flow 0.25 mL min^{-1}, column temperature 40°C, time 0 min, 5% B, time 20 min, 100% B.

A SpectraSystem UV6000LP detector (Thermo Separation Products) measured UV/Vis absorption between 200 and 450 nm with a scan rate of 2 scans/s. Atmospheric pressure chemical ionization, operated in the negative ionization mode, was used as an ion source to couple HPLC with an MS instrument (LCQ Classic; ThermoQuest, San Jose, CA; vaporizer temperature 450°C, capillary temperature 150°C, source current 5 μA, sheath gas flow 21, aux gas flow 3).

During separation, the most abundant ion in each full MS scan was fragmented in the next scan with the dependent MS/MS mode.

Additionally, each fraction was separated on LC-MS/MS under higher acidity buffer conditions. A gradient separation was run from solvent C (1% aqueous acetic acid, pH 2) to solvent D (acetonitrile, 1% acetic acid) under the following conditions: column temperature 40°C, flow 0.3 mL min^{-1}, time 0 min, 5% D, time 1 min, 17% D, time 19 min, 77% D, time 20 min, 100% D. The MS conditions were: vaporizer temperature 350°C, capillary temperature 100°C, source current 5 μA, sheath gas flow set at 34, aux gas flow set at 4.

NMR Spectroscopy

Compounds were authenticated by the normal range of 1D and standard 2D (COSY, TOCSY, HSQC, HMBC) experiments on a 360 MHz DRX-360 instrument (Bruker, Karlsruhe, Germany) fitted with a 5-mm ^{1}H/broadband gradient probe with inverse geometry (proton coils closest to the sample). The solvent was acetone-d_6 unless otherwise noted; the central acetone solvent peak was used as internal reference (δ_C 29.8, δ_H 2.04 ppm). NMR data will be deposited in our NMR database of lignin and cell wall model compounds (http://www.dfrc.ars.usda.gov/software.html; Ralph et al., 2004a).

Shorthand Naming Convention for Oligolignols (Dimers, Trimers, and Tetramers)

To describe the oligolignols in a logical and informative manner, the following convention has been adopted. Bold **G** and **S** are used for guaiacyl and syringyl units, to name the units derived from coupling reactions of coniferyl and sinapyl alcohol; bold **SP** for units derived from the incorporation of sinapyl *p*-hydroxybenzoate esters; and **G′**, **S′**, and **V′** for units derived from coniferaldehyde, sinapaldehyde, and vanillin, respectively. The interunit bond formed during the radical coupling reaction is specified in parentheses: (8–O–4), (8–5), or (8–8). For example, **G**(8–O–4)**S**(8–5)**G′** results from sinapyl alcohol coupling at its C_8 position with coniferaldehyde at its C_5 position to make dimer **S**(8–5)**G′**, followed by coupling of this dimer at its phenolic 4–O position with another coniferyl alcohol radical at its favored C_8 position. The descriptor for the trimer is unambiguous because coupling the other way round is not possible because coupling always requires a free-phenolic group on the unit's aromatic ring; for instance, first coupling of coniferyl alcohol at its C_8 position with sinapyl alcohol at its 4–O position could produce the dimer **G**(8–O–4)**S**, but the specified trimer can no longer result from further coupling to this dimer, because only the **G** unit in the dimer is capable of entering coupling reactions (Ralph et al., 2004b). Whenever it could be determined or tentatively identified, *erythro*- and *threo*-isomers of the (8–O–4) structures are indicated as (*e*8–O–4) and (*t*8–O–4). The shorthand notation (8–O–4*) indicates a benzylic oxidized (8–O–4)-linked unit, i.e. bearing a 7-oxo group (Fig. 4, compound **30**).

Preparation and HPLC Analysis of Oligolignol Mixtures

G, **S**, and **G** + **S** synthetic oligolignol mixtures were prepared by using $Cu(OAc)_2$ oxidation of coniferyl alcohol, sinapyl alcohol, or both coniferyl and sinapyl alcohols, respectively, as described previously (Landucci et al., 1995). For example, the **S** + **G** oligolignols were prepared as follows. A mixture of coniferyl alcohol (480 mg, 2.66 mmol) and sinapyl alcohol (560 mg, 2.66 mmol) was dissolved in 0.05 M borate solution (200 mL, pH 9.2), which was prepared from sodium borate ($Na_2B_4O_7{\cdot}10H_2O$) at 100°C with stirring. A solution of copper acetate (1.06 g, 5.32 mmol) in water (10 mL) was added into the reaction solution. A green suspension formed and was stirred for 30 min at 100°C, after which time the solution color turned to orange. Insoluble copper salts were removed by filtration and the bright yellow solution was collected. The solution was acidified to pH 5 with 0.5 M HCl and extracted with ethyl acetate. The yellow extract was dried over $MgSO_4$ and concentrated by rotary evaporation. The boric acid was removed as methyl borate by adding MeOH and evaporating three times. A white foamy solid (83%) was obtained. Part of the crude product was used to separate compounds by thin-layer chromatography (TLC; $CHCl_3$:MeOH; 10:1) to isolate and identify the structures. The mixtures were also used for HPLC analysis under conditions comparable to those used for the poplar methanol extracts. Coniferyl alcohol (200 mg) or sinapyl alcohol (250 mg) were treated independently in a similar manner to prepare oligolignols consisting of **G** or **S** units only. Higher amounts of some compounds were obtained in $Mn(OAc)_3$/pyridine reactions (Landucci et al., 1995).

The following compounds were purified from these preparations and identified by MS and NMR analysis: **G**(*t*8–O–4)**G**, **1**; **S**(8–5)**G**, **6**; **S**(8–O–4)**S′**, **9**; **S**(*t*8–O–4)**S**(8–8)**S**, **16**; **S**(*t*8–O–4)**S**(8–5)**G**, **17**; **G**(*t*8–O–4)**S**(8–5)**G**, **18**; **G**(*e*8–O–4)**S**(8–5)**G**, **20**; **G**(*t*8–O–4)**S**(8–8)**G**, **22**; **G**(*t*8–O–4)**S**(8–5)**G′**, **23**; **G**(*t*8–O–4)**S**(8–8)**S**, **24**; **G**(*e*8–O–4)**S**(8–5)**G′**, **25**; **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, **33**.

Chemical Synthesis of Dilignols, Trilignols, and Tetralignols

To authenticate the compounds isolated from the xylem fraction, the required compounds were prepared by the above $Cu(OAc)_2$ oxidation unless described specifically. Non-HPLC separations were by preparative TLC or by flash chromatography. Where available, the compound number in our NMR database of lignin and cell wall model compounds (http://www.dfrc.ars.usda.gov/software.html; Ralph et al., 2004a) is noted.

Compound **1**, **G**(*t*8–O–4)**G**, *threo*-1-(4-hydroxy-3-methoxy-phenyl)-2-[4-(3-hydroxy-propenyl)-2-methoxy-phenoxy]-propane-1,3-diol, (8–O–4)-dehydrodiconiferyl alcohol), database number 2013*t*: prepared from the $Cu(OAc)_2$ system described above using coniferyl alcohol. Oxidation of coniferyl alcohol by $Fe(NH_4)(SO_4)_3$ gave a mixture of *threo*- and *erythro*-isomers.

Compound **2**, **G**(8–5)**G**, 4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydrobenzofuran-2-yl]-2-methoxy-phenol], (8–5)-dehydrodiconiferyl alcohol, database number 2004: prepared as described previously (Quideau and Ralph, 1994).

Compound **3**, **G**(8–8)**G**, 1,4-bis-(4-hydroxy-3-methoxy-phenyl)-tetrahydrofuro[3,4-c]furan], pinoresinol, database number 2020: prepared according to Syrjänen and Brunow (2000).

Compound **4**, **S**(8–8)**S**, 1,4-bis-(3,5-dimethoxy-4-hydroxy-phenyl)-tetrahydro-furo[3,4-c]furan], syringaresinol, database number 117: synthesized by coupling sinapyl alcohol via $CuSO_4$, as described previously (Freudenberg et al., 1958).

Compound **6**, **S**(8–5)**G**, 4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydrobenzofuran-2-yl]-2,6-dimethoxy-phenol, simulanol, database number 3063: prepared from the $Cu(OAc)_2$ system described above using coniferyl and sinapyl alcohols.

Compound **7**, **G**(8–5)**G′**, 3-[2-(4-hydroxy-3-methoxy-phenyl)-3-hydroxymethyl-7-methoxy-2,3-dihydro-benzofuran-5-yl]-propenal, balanophonin, database number 2021, is the cinnamaldehyde analog of compound **G**(8–5)**G**, **2**. For its synthesis, compound **2** (70 mg, 0.21 mmol) was dissolved in tetrahydrofuran (10 mL) and 2,3-dichloro-5,6-dicyanobenzoquinone (Becker et al., 1980) was added. The reaction mixture was stirred overnight at room temperature. Crude products were dried and compound **7** **G**(8–5)**G′** was isolated by TLC (chloroform:ethyl acetate; 1:1). Yellow needle crystals (55.4 mg, yield 74%) were obtained from acetone-petroleum ether. NMR data were consistent with those reported previously (Quideau and Ralph, 1994; Sy and Brown, 1999).

Compound **8**, **G**(*t*8–O–4)**S**, *threo*-1-(4-hydroxy-3-methoxy-phenyl)-2-[2,6-dimethoxy-phenoxy-4-(3-hydroxy-propenyl)]-propane-1,3-diol, database number 3067: small amounts were reported (Landucci et al., 1995) in the oligomer reactions described above, but were insufficient to isolate here. The peracetate (database number 188) was previously isolated, following acetylation, from $Mn(OAc)_3$ reactions in pyridine (Landucci et al., 1995). Here, compound **8** was more conveniently isolated from synthetic oligolignol mixtures from $Cu(OAc)_2$ oxidations of coniferyl plus sinapyl alcohols in acetone at room temperature.

Compound **9**, **S**(8–O–4)**S′**, (2*E*)-3-{4-[2-hydroxy-1-(hydroxymethyl)-2-(4-hydroxy-3,5-dimethoxyphenyl)-ethoxy]-3,5-dimethoxyphenyl-acrylaldehyde: prepared from the $Mn(OAc)_3$ oxidation of sinapyl alcohol in pyridine (Landucci et al., 1995).

Compound **11**, **G**(8–5)**V′**, 2-(4-hydroxy-3-methoxy-phenyl)-3-hydroxymethyl-7-methoxy-2,3-dihydro-benzofuran-5-carbaldehyde, ficusal, database number 3061: prepared in low yields from **G**(8–5)**G** **2** by oxidation with CrO_3/montmorillonite K-10 in CH_2Cl_2 (which also produced low yields of compound **7**; Heravi et al., 1999).

Compound **16**, **S**(*t*8–O–4)**S**(8–8)**S**, 1-(4-hydroxy-3,5-dimethoxy-phenyl)-2-{4-[4-(4-hydroxy-3,5-dimethoxy-phenyl)-tetrahydro-furo[3,4-*c*]furan-1-yl]-2,6-dimethoxy-phenoxy}-propane-1,3-diol, *threo*-buddlenol D, database number 198: isolated from the oligomers of the above $Cu(OAc)_2$ oxidation of sinapyl alcohol.

Compound **17**, **S**(*t*8–O–4)**S**(8–5)**G**, 1-(4-hydroxy-3,5-dimethoxy-phenyl)-2-{4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydro-benzofuran-2-yl]-2,6-dimethoxy-phenoxy}-propane-1,3-diol: present in oligolignol mixtures from $Cu(OAc)_2$ oxidations of coniferyl plus sinapyl alcohols described above.

Compounds **18** and **20**, **G**(*t*8–O–4)**S**(8–5)**G** and **G**(*e*8–O–4)**S**(8–5)**G**, *threo*- and *erythro*-1-(4-hydroxy-3-methoxy-phenyl)-2-{4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydro-benzofuran-2-yl]-2,6-dimethoxy-phenoxy}-propane-1,3-diol, buddlenol B, database number 181: prepared from the $Cu(OAc)_2$ system described above using coniferyl and sinapyl alcohols (Landucci et al., 1995). NMR showed that the synthesized compound was mainly present in the *threo* form.

Compound **19**, **SP**(8–8)**S**, tetrahydro-α4,2-*bis*-(4-hydroxy-3,5-dimethoxyphenyl)α3-O-(4-hydroxybenzoyl)-3,4-furandimethanol, database number 3066: this compound, crucial to establishing **SP** as a lignin precursor, was prepared by coupling sinapyl alcohol and sinapyl *p*-hydroxybenzoate via peroxidase/H_2O_2 (approximately 30% yield) as described in detail elsewhere (Lu et al., 2004).

Compound **22**, **G**(*t*8–O–4)**S**(8–8)**G**, 1-(4-hydroxy-3-methoxy-phenyl)-2-{4-[4-(4-hydroxy-3-methoxy-phenyl)-tetrahydro-furo[3,4-c]furan-1-yl]-2,6-dimethoxy-phenoxy}-propane-1,3-diol, buddlenol E, database number 3064: isolated from the synthetic oligolignol mixture from $Cu(OAc)_2$ oxidation of sinapyl and coniferyl alcohols described above.

Compounds **23** and **25**, **G**(*t*8–O–4)**S**(8–5)**G′** and **G**(*e*8–O–4)**S**(8–5)**G′**, *threo*- and *erythro*-3-(2-{4-[2-hydroxy-2-(4-hydroxy-3-methoxy-phenyl)1-hydroxymethyl-ethoxy]-3,5-dimethoxy-phenyl}-3-hydroxymethyl-7-methoxy-2,3-dihydro-benzofuran-5-yl)-propenal, buddlenol A, database number 3065: isolated from the synthetic oligolignol mixture from $Cu(OAc)_2$ oxidation of sinapyl and coniferyl alcohols described above. NMR showed that the synthesized compound was mainly present in the *threo* form.

Compound **24**, **G**(*t*8–O–4)**S**(8–8)**S**, 2-{4-[4-(4-hydroxy-3,5-dimethoxy-phenyl)-tetrahydrofuro[3,4-c]furan1-yl]-2,6-dimethoxy-phenoxy}-1-(4-hydroxy-3-methoxy-phenyl)-propane-1,3-diol, buddlenol C, database number 183: isolated from the synthetic oligolignol mixture from $Cu(OAc)_2$ oxidation of sinapyl and coniferyl alcohols described above.

Compound **33**, **G**(8–O–4)**S**(8−8)**S**(8–O–4)**G**, 2-[4-(4-{4-[2-hydroxy-2-(4-hydroxy-3-methoxy-phenyl)1-hydroxymethylethoxy]-3,5-dimethoxy-phenyl}-tetrahydro-furo[3,4-c]furan-1-yl)-2,6-dimethoxy-phenoxy]-1-(4-hydroxy-3-methoxy-phenyl)-propane-1,3-diol, hedyotisol, database number 194: isolated from the synthetic oligolignol mixture from $Cu(OAc)_2$ oxidation of sinapyl and coniferyl alcohols described above. The *threo/erythro* configurations of the two β-aryl ether units were not determined.

Compound **39**, **S**(*t*8–O–4)**G**(8–5)**G**, 1-(4-hydroxy-3,5-dimethoxy-phenyl)-2-{4-[3-hydroxymethyl-5-(3-hydroxy-propenyl)-7-methoxy-2,3-dihydro-benzofuran-2-yl]-2-methoxy-phenoxy}-propane-1,3-diol, was prepared via traditional synthetic β-ether lignin model methods, for instance as in Ralph et al. (1986). Briefly, the coniferyl alcohol dimer **G**(8–5)**G** (compound **2**, (8–5)-dehydrodiconiferyl alcohol) was added to acetate-protected α-bromo-acetosyringone, formaldehyde was added (to create the three-carbon side chain) and the benzylic ketone was reduced with $NaBH_4$ in ethanol:H_2O (1:1). Completing the deacetylation in pyrrolidine:MeOH (1:1) provided **S**(*t*8–O–4)**G**(8–5)**G**, **39**. NMR (data from 2D, average chemical shift values only, using **A**(*t*8–O–4)**B**(8–5)**C** to unambiguously identify the units in **S**(*t*8–O–4)**G**(8–5)**G**): δ_C/δ_H 105.3/6.76 (**A**2/6), 130.4/6.53 (**C**7), 128.3/6.23 (**C**8), 88.0/5.58 (**B**7), 73.8/4.87 (**A**7), 87.9/4.24 and 86.2/4.33 (**A**8), 62.2/4.19 (**C**9), 64.6/3.88,3.82 (**B**9), 61.9/3.8-3.7,3.51 (**A**9), 56.5/3.78 and 56.3/3.86 (OMe), 54.8/3.52 (**B**8).

Distribution of Materials

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining any permissions will be the responsibility of the requester.

ACKNOWLEDGMENT

The authors thank Martine De Cock for help in preparing the manuscript.

Received July 7, 2004; returned for revision September 28, 2004; accepted September 28, 2004.

LITERATURE CITED

Adler E (1977) Lignin chemistry - past, present and future. Wood Sci Technol **11:** 169–218

Akiyama T, Matsumoto Y, Okuyama T, Meshitsuka G (2003) Ratio of *erythro* and *threo* forms of β-O-4 structures in tension wood lignin. Phytochemistry **64:** 1157–1162

Baucher M, Chabbert B, Pilate G, Van Doorsselaere J, Tollier M-T, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inzé D, et al (1996) Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. Plant Physiol **112:** 1479–1490

Baucher M, Halpin C, Petit-Conil M, Boerjan W (2003) Lignin: genetic engineering and impact on pulping. Crit Rev Biochem Mol Biol **38:** 305–350

Baucher M, Monties B, Van Montagu M, Boerjan W (1998) Biosynthesis and genetic engineering of lignin. Crit Rev Plant Sci **17:** 125–197

Becker HD, Bjoerk A, Adler E (1980) Quinone dehydrogenation. Oxidation of benzylic alcohols with 2,3-dichloro-5,6-dicyanobenzoquinone. J Org Chem **45:** 1596–1600

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol **54:** 519–546

Boudet AM, Kajita S, Grima-Pettenati J, Goffner D (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. Trends Plant Sci **8:** 576–581

Brunow G, Karlsson O, Lundquist K, Sipilä J (1993) On the distribution of the diastereomers of the structural elements in lignins: the steric course of reactions mimicking lignin biosynthesis. Wood Sci Technol **27:** 281–286

Brunow G, Lundquist K, Gellerstedt G (1999) Lignin. *In* E Sjöström, R Alén, eds, Analytical Methods in Wood Chemistry, Pulping and Papermaking. Springer-Verlag, Berlin, pp 77–124

Burlat V, Kwon M, Davin LB, Lewis NG (2001) Dirigent proteins and dirigent sites in lignifying tissues. Phytochemistry **57:** 883–897

Chen Y-R, Sarkanen S (2004) Macromolecular lignin replication: a mechanistic working hypothesis. Phytochemistry Rev (in press)

Christensen JH, Baucher M, O'Connell AP, Van Montagu M, Boerjan W (2000) Control of lignin biosynthesis *In* SM Jain, SC Minocha, eds, Molecular Biology of Woody Plants, Vol 1 (Forestry Sciences, Vol 64). Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 227–237

Davin LB, Wang H-B, Crowell AL, Bedgar DL, Martin DM, Sarkanen S, Lewis NG (1997) Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. Science **275:** 362–366

Faix O (1986) Investigation of lignin polymer models (DHP's) by FTIR spectroscopy. Holzforschung **40:** 273–280

Freudenberg K (1959) Biosynthesis and constitution of lignin. Nature **183:** 1152–1155

Freudenberg K, Harkin JM, Reichert M, Fukuzumi T (1958) Die Dehydrierung des Sinapinalkohols. Chem Ber **91:** 581–590

Freudenberg K, Neish AC (1968) Constitution and Biosynthesis of Lignin. Springer-Verlag, Berlin

Grabber JH, Ralph J, Hatfield RD, Quideau S, Kuster T, Pell AN (1996) Dehydrogenation polymer—cell wall complexes as a model for lignified grass walls. J Agric Food Chem **44:** 1453–1459

Guo D, Chen F, Wheeler J, Winder J, Selman S, Peterson M, Dixon RA (2001) Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin *O*-methyltransferases. Transgenic Res **10:** 457–464

Hatfield R, Vermerris W (2001) Lignin formation in plants. The dilemma of linkage specificity. Plant Physiol **126:** 1351–1357

Heravi MM, Kiakojoori R, Tabar-Hydar K (1999) Ammonium chlorochromate adsorbed on montmorillonite K-10: a new reagent for the oxidation of alcohols to the corresponding carbonyl compounds under non-aqueous conditions. Monatsh Chem **130:** 581–583

Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlén M, Teeri TT, Lundeberg J, et al (2001) A transcriptional roadmap to wood formation. Proc Natl Acad Sci USA **98:** 14732–14737

Kim H, Ralph J, Lu F, Ralph SA, Boudet A-M, MacKay JJ, Sederoff RR, Ito T, Kawai S, Ohashi H, et al (2003) NMR analysis of lignins in CAD-deficient plants. Part 1. Incorporation of hydroxycinnamaldehydes and hydroxybenzaldehydes into lignins. Org Biomol Chem **1:** 268–281

Landucci LL, Deka GC, Roy DN (1992) A ^{13}C NMR study of milled wood lignins from hybrid *Salix* clones. Holzforschung **46:** 505–511

Landucci LL, Luque S, Ralph J (1995) Reaction of *p*-hydroxycinnamyl

alcohols with transition metal salts. 2. Preparation of guaiacyl/syringyl di-, tri- and tetralignols. J Wood Chem Technol **15:** 493–513

Landucci LL, Ralph SA (2001) Reactions of *p*-hydroxycinnamyl alcohols with transition metal salts. IV. Tailored syntheses of β-O-4 trimers. J Wood Chem Technol **21:** 31–52

Lewis NG (1999) A 20th century roller coaster ride: a short account of lignification. Curr Opin Plant Biol **2:** 153–162

Lewis NG, Davin LB (1999) Lignans: biosynthesis and function. *In* U Sankawa, ed, Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives, Comprehensive Natural Products Chemistry, Vol 1. Elsevier Science, Amsterdam, pp 639–712

Lu F, Ralph J (1997) DFRC method for lignin analysis. 1. New method for β-aryl ether cleavage: lignin model studies. J Agric Food Chem **45:** 4655–4660

Lu F, Ralph J (2002) Preliminary evidence for sinapyl acetate as a lignin monomer in kenaf. Chem Commun **2002:** 90–91

Lu F, Ralph J (2003) Non-degradative dissolution and acetylation of ball-milled plant cell walls: high-resolution solution-state NMR. Plant J **35:** 535–544

Lu F, Ralph J, Morreel K, Messens E, Boerjan W (2004) Preparation and relevance of a cross-coupling product between sinapyl alcohol and sinapyl p-hydroxybenzoate. Org Biomol Chem **2:** 2888–2890

Lundquist K (1992) Acidolysis. *In* SY Lin, CW Dence, eds, Methods in Lignin Chemistry. Springer-Verlag, Heidelberg, pp 287–300

Meyermans H, Morreel K, Lapierre C, Pollet B, De Bruyn A, Busson R, Herdewijn P, Devreese B, Van Beeumen J, Marita JM, et al (2000) Modification in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. J Biol Chem **275:** 36899–36909

Morreel K, Ralph J, Lu f, Goeminne G, Busson R, Herdewijn P, Goeman J, Van der Eycken j, Boerjan W, Messens E (2004) Phenolic profiling of caffeic acid *O*-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols. Plant Physiol **136:** (in press)

Nakano J, Ishizu A, Migata N (1961) Studies on lignin. XXXII. Ester groups of lignin. Tappi **44:** 30–32

Pilate G, Guiney E, Holt K, Petit-Conil M, Lapierre C, Leplé J-C, Pollet B, Mila I, Webster EA, Marstorp HG, et al (2002) Field and pulping performances of transgenic trees with altered lignification. Nat Biotechnol **20:** 607–612

Quideau S, Ralph J (1994) A biomimetic route to lignin model compounds *via* silver (I) oxide oxidation. 1. Synthesis of dilignols and non-cyclic benzyl aryl ethers. Holzforschung **48:** 12–22

Ralph J, Ede RM, Wilkins AL (1986) Synthesis of trimeric lignin model compounds composed of β-aryl ether and phenylcoumaran structures. Holzforschung **40:** 23–30

Ralph SA, Landucci LL, Ralph J (2004a) NMR database of lignin and cell wall model compounds. US Dairy Forage Research Center, US Department of Agriculture, Agricultural Research Service. http://www.dfrc.ars.usda.gov/software.html (November 15, 2004)

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, et al (2004b) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochemistry Rev **3:** (in press)

Ralph J, Marita JM, Ralph SA, Hatfield RD, Lu F, Ede RM, Peng J, Quideau S, Helm RF, Grabber JH, Kim H, et al (1999a) Solution-state NMR of lignins. *In* DS Argyropoulos, ed, Advances in Lignocellulosic Characterization. TAPPI Press, Atlanta, pp 55–108

Ralph J, Peng J, Lu F, Hatfield RD, Helm RF (1999b) Are lignins optically active? J Agric Food Chem **47:** 2991–2996

Rolando C, Monties B, Lapierre C (1992) Thioacidolysis. *In* SY Lin, CW Dence, eds, Methods in Lignin Chemistry. Springer-Verlag, Berlin, pp 334–349

Sakakibara N, Suzuki S, Umezawa T, Shimada M (2003) Biosynthesis of yatein in *Anthriscus sylvestris*. Org Biomol Chem **1:** 2474–2485

Sarkanen KV (1971) Precursors and their polymerization. *In* KV Sarkanen, CH Ludwig, eds, Lignins, Occurrence, Formation, Structure and Reactions. Wiley-Interscience, New York, pp 95–163

Smith DCC (1955) *p*-Hydroxybenzoate groups in the lignin of aspen (*Populus tremula*). J Chem Soc **1955:** 2347–2351

Snyder LR, Kirkland JJ, Glajch JL (1997) Practical HPLC Method Development, Ed 2. John Wiley & Sons, New York

Sun RC, Fang JM, Tomkinson J (1999) Fractional isolation and structural characterization of lignins from oil palm trunk and empty fruit bunch fibres. J Wood Chem Technol **19:** 335–356

Sy L-K, Brown GD (1999) Coniferaldehyde derivatives from tissue culture of *Artemisia annua* and *Tanacetum parthenium*. Phytochemistry **50:** 781–785

Syrjänen K, Brunow G (1998) Oxidative cross coupling of *p*-hydroxycinnamic alcohols with dimeric arylglycerol β-aryl ether lignin model compounds. The effect of oxidation potentials. J Chem Soc [Perkin 1] **1998:** 3425–3429

Syrjänen K, Brunow G (2000) Regioselectivity in lignin biosynthesis. The influence of dimerization and cross-coupling. J Chem Soc [Perkin 1] **2000:** 183–187

Umezawa T (2004) Diversity of lignan biosynthesis. Phytochemistry Rev **3:** (in press)

Zhong R, Morrison WH III, Himmelsbach DS, Poole FL II, Ye Z-H (2000) Essential role of caffeoyl coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. Plant Physiol **124:** 563–577

SUPPLEMENTARY MATERIAL

Oligolignol structure elucidation

G-type dimers

*Compound **1***, **G**(*t*8–O–4)**G**

This compound eluted at 11.5 min and its UV/VIS spectrum showed maximum absorptions at 203.0 and 265.3 nm. MS/MS analysis indicated a molecular weight (MW) of 376 g/mol. The presence of two daughter ions at *m/z* 179 and 195 (See Fig. Supp. Mat.), suggested a possible structure whereby two **G** units are connected by an 8–O–4-bond. Based on the relative abundance of each fragment ion, the collision-activated cleavage of this bond prefers the release of a neutral coniferyl alcohol fragment (180 Da) rather than a neutral hydroxylated coniferyl alcohol residue (196 Da). Neutral losses led to daughter ions at *m/z* 357 (loss of H_2O), 345 (loss of CH_2O) and 327 (combined loss of H_2O and CH_2O); this pattern of peaks was only found in the MS/MS spectra of β-aryl ether unit-containing oligolignols (see below). Because the ion at *m/z* 327 was much more prominent than the ions at *m/z* 357 and 345, this suggested a *threo*-configuration about the 8–O–4-bond (see MS/MS spectra of *threo*- and *erythro*-buddlenol B, compounds **18** and **20**). Hence, **G**(*t*8–O–4)**G**, *threo*-(8–O–4)-dehydrodiconiferyl alcohol (Lewis and Davin, 1999) was synthesized. The synthetic compound had an identical retention time and MS/MS spectrum as that of the natural isolated compound **1**.

*Compound **2***, **G**(8–5)**G**

This peak was characterized by a retention time of 15.1 min and absorption maxima at 225.3 and 275.9 nm in the UV/VIS spectrum. The MS/MS data indicated a molecular weight of 358 g/mol. A dilignol composed of two **G** units, connected by an 8–5-linkage (i.e. a phenylcoumaran unit) or an 8–8-linkage (i.e. a resinol unit) was proposed. Both compounds were synthesized, but only the phenylcoumaran **G**(8–5)**G**, (8–5)-dehydrodiconiferyl alcohol (Lewis and Davin, 1999), showed the same retention time and MS/MS spectrum as the natural compound.

*Compound **3***, **G**(8–8)**G**

The UV/VIS spectrum (λ_{max} at 230.0 and 279.4 nm) of this compound, eluting at 17.3 min, was neither similar to that of **G**(*t*8–O–4)**G**, nor to that of the phenylcoumaran **2**, yet its molecular weight of 358 g/mol suggested two **G** units. In addition to β-aryl ether and phenylcoumaran units, resinol units are major bonding structures formed during dehydrodimerization of monolignol radicals (Adler, 1977; Boerjan et al., 2003). So, pinoresinol (Umezawa, 2004), **G**(8–8)**G**, was proposed for this compound. Spiking of synthesized pinoresinol authenticated this structure.

S-type dimers

*Compound **4***, **S**(8–8)**S**

The UV/VIS spectrum of this compound, eluting at 16.6 min, showed absorption maxima at 218.3 and 271.2 nm. Its MS/MS spectrum (See Fig. Supp. Mat.) indicated a molecular weight of 418 g/mol, which suggested the presence of two **S** units. The MS/MS spectral pattern was similar to that of pinoresinol, **G**(8–8)**G**, but diagnostic *m/z* values were shifted by 30 or 60 g/mol, indicating that this compound is syringaresinol (Umezawa, 2004), **S**(8–8)**S**. Spiking with synthetic syringaresinol confirmed this.

*Compound **5***, **S**(*t*8–O–4)**S**

LC-MS analysis of fraction 6 (Fig. 3), isolated by HPLC, showed the presence of a compound of molecular mass 436 g/mol. The intense fragment ion at *m/z* 387, and fragment ions at *m/z* 405 and 417 indicated a β-aryl ether unit, the *threo*-isomer being consistent with what is found for compounds **18** and **20** with their known *threo*- and *erythro*-configurations. Cleaving the β-aryl ether unit presumably generates the ions at *m/z* 209 and 225. So, the structure **S**(*t*8–O–4)**S**, i.e. *threo*-(8–O–4)-dehydrodisinapyl alcohol (Cutillo et al., 2003), was indicated by these

MS/MS data.

Mixed-type dimers

*Compound **6***, **S**(8–5)**G**

The UV/VIS spectrum (λ_{max} at 219.5 and 273.5 nm) of the peak at 14.8 min was similar to that of phenylcoumaran **G**(8–5)**G**, **2**. A molecular weight of 388 g/mol was obtained for this compound by LC-MS/MS (See Fig. Supp. Mat.), 30 g/mol more than **G**(8–5)**G**. The same neutral losses were found as in the MS/MS spectrum of **G**(8–5)**G**. Hence, **S**(8–5)**G** or simulanol (Yang et al., 2002) was proposed as the structure of this compound. The synthetic compound confirmed the proposed structure. Fragment ions at *m/z* 221, 203 and 191 were also found in the spectrum of **G**(8–5)**G** and can be attributed to the **G** unit bearing the propenol side-chain.

Compound 7, **G**(8–5)**G'**

The retention time of this product was 17.5 min and the UV/VIS spectrum showed absorption maxima at 230.0 and 345.8 nm. A molecular weight of 356 g/mol was obtained by means of LC-MS/MS. The neutral losses corresponding with the two most abundant daughter ions (*m/z* 219 and 337) were the same as in the MS/MS spectrum of (8–5)-dehydrodiconiferyl alcohol, **G**(8–5)**G**, **2**. The low-abundant ion at *m/z* 327 originates probably from loss of CO, suggesting an aldehyde. Based on these data, this compound was suspected to be balanophonin **G**(8–5)**G'** (Haruna et al., 1982), the cinnamaldehyde analog of (8–5)-dehydrodiconiferyl alcohol **G**(8–5)**G**, produced either via post-coupling oxidation of the cinnamyl alcohol, or via coupling of coniferyl alcohol (at its C_8-position) with coniferaldehyde (at its 4–O-position). The synthetic compound confirmed the proposed structure.

*Compound **8***, **G**(*t*8–O–4)**S**

A peak characterized by a molecular weight of 406 g/mol was detected in fraction 5 (Fig. 3) by means of LC-MS/MS. Both the molecular weight and MS/MS daughter ions at *m/z* 357 (base peak), 375 and 387 suggested the β-aryl ether analog of compound **1** involving an **S** and a **G** unit. This was further verified by the presence of MS/MS fragments at *m/z* 195 and 209, indicating that the structure is **G**(8–O–4)**S**, the *threo*-isomer being consistent with the data, i.e. **G**(*t*8–O–4)**S** or *threo*-guaiacylglycerol 8–O–4 sinapyl ether (Lewis & Davin, 1999). This structure was authenticated by MS/MS analysis of the synthesized compound.

*Compound **9***, **S**(8–O–4)**S'**

LC-MS/MS analysis of fraction 12 revealed the presence of a compound possessing a molecular weight of 434 g/mol. This suggested that the structure was the corresponding cinnamaldehyde of **S**(*t*8–O–4)**S**, **5**. This structure was synthesized and its MS/MS spectrum, dominated by neutral losses of 30 (*m/z* 403) and 60 g/mol (*m/z* 373), matched that of the natural compound. However, the biological and the synthesized compound **9** are not the same stereomers since the synthesized compound eluted at a different time than the isolated fraction in which the biological compound was found (data not shown).

*Compound **10***, **G**(8–O–4)**S'** or **S**(8–O–4)**G'**

A compound with a molecular weight of 404 g/mol was found by the LC-MS/MS separation of fraction 15. The MS/MS spectrum showed the same neutral losses as observed in the spectrum of the previous compound, **S**(8–O–4)**S'**, **9**. Hence, this compound is **G**(8–O–4)**S'** or **S**(8–O–4)**G'**.

*Compound **11***, **G**(8–5)**V'**

LC-MS/MS analysis of fraction 15 revealed a compound with a molecular weight of 330 g/mol. MS/MS daughter ions at *m/z* 193, 178 and 164 were reminiscent of a phenylcoumaran containing an aromatic aldehyde when compared with the pattern of peaks at *m/z* 219, 204 and 190 in the MS/MS spectrum of compound **7**, **G**(8–5)**G'**. The neutral loss corresponding with the daughter ion at *m/z* 193 was indicative of a **G** unit connected to a unit derived

from vanillin, i.e. **G**(8–5)**V'** or ficusal (Li & Kuo, 2000), a structure which was finally authenticated by MS/MS analysis of the synthesized compound.

*Compound **12***, **S**(8–5)**G'**

This compound (MW = 386 g/mol) was found by LC-MS analysis of fraction 21. The MS/MS spectrum of this compound was similar to that of **G**(8–5)**G'**, **7**. Daughter ions at *m/z* 219, 204 and 190, characteristic of a phenylcoumaran involving a unit derived from coniferaldehyde, showed that this compound is the aldehyde analogue of compound **6**, **S**(8–5)**G**, i.e. **S**(8–5)**G'**.

*Compound **13***, **G**(8–O–4)**G'**

A compound with a molecular weight of 374 g/mol was detected in fraction 22. The structure of this compound was resolved as **G**(8–O–4)**G'**, since the principal MS/MS fragmentations were associated with neutral losses of 30 and 60 g/mol as was also found in the spectra of compounds **9** and **10**, i.e. **S**(8–O–4)**S'** and **G**(8–O–4)**S'** or **S**(8–O–4)**G'**.

G-type trimers

*Compound **14***, **G**(*t*8–O–4)**G**(*t*8–O–4)**G**

At 11.8 min, a compound eluted with a UV/VIS spectrum (λ_{max} at 201.9 and 266.4 nm) similar to **G**(*t*8–O–4)**G** and a molecular weight of 572 g/mol. This suggested the presence of three **G** units connected by β-aryl ether bonds, which was substantiated by the detection of MS/MS daughter ions at *m/z* 523, 541, and 553, owing to the elimination of neutral CH_2O and/or H_2O from the parent ion. The cleavage of one of the β-aryl ether units yields the ions at *m/z* 195 and 375, whereas the ions at *m/z* 179 and 391 are associated with the fragmentation of the other β-aryl ether unit. Elimination of CH_2O/H_2O from the fragments *m/z* 375 and 391 yields ions at *m/z* 327 and 343. Together, these results indicate a **G**(8–O–4)**G**(8–O–4)**G** structure tentatively with both β-ether units in their *threo*-forms, i.e. **G**(*t*8–O–4)**G**(*t*8–O–4)**G**.

*Compound **15***, **G**(*t*8–O–4)**G**(8–8)**G**

Using LC-MS, a compound was found in fraction 18 (Fig. 3) having a molecular weight of 554 g/mol. This suggested a possible trimer involving a β-aryl ether unit, because daughter ions were found at *m/z* 535, 523 and 505 (base peak) (see compound **1**). The cleavage of this bond produced daughter ions associated with each fragment, i.e. at *m/z* 195 and 357. The latter ion is not observed in the MS/MS spectra of trimers involving both a phenylcoumaran and a β-aryl ether unit (see below). So, these data are in favor of a **G** unit 8–O–4-linked to a **G**(8–8)**G** substructure, i.e. **G**(8–O–4)**G**(8–8)**G**, tentatively with a *threo*-β-ether unit, i.e. **G**(*t*8–O–4)**G**(8–8)**G**.

S-type trimers

*Compound **16***, **S**(*t*8–O–4)**S**(8–8)**S**

At 17.4 min, a trilignol (MW = 644 g/mol) only composed of **S** units eluted. The UV/VIS spectrum showed maximal absorptions at 206.6 and 271.2 nm. A *threo*-β-aryl ether unit is indicated by the presence and ratio of product ions at *m/z* 595, 613 and 625 in the MS/MS spectrum. An **S**(8–O–4) unit is indicated by the fragment *m/z* 225, attached to a syringaresinol unit (*m/z* 417). This compound, **S**(*t*8–O–4)**S**(8–8)**S**, buddlenol D (Houghton, 1985), was purified from the synthetic oligolignol mixture from $Cu(OAc)_2$ oxidation of sinapyl alcohol as described in the material and methods and used for the authentication of the natural compound.

Mixed-type trimers

*Compound **17***, **S**(*t*8–O–4)**S**(8–5)**G**

The UV/VIS spectrum (λ_{max} at 213.6 and 275.9 nm) of the compound eluting at 15.8 min was similar to those of phenylcoumarans. Furthermore, collision-activation of this compound (MW = 614 g/mol) generated product ions indicating the presence of an **S** unit (*m/z* 225) attached by a *threo*-β-aryl ether unit (*m/z* 565, 583 and 595) to an **S-G** dimeric substructure. The fragments *m/z* 354, 357 and 369 were also found in about the same relative abundances in the MS/MS spectrum of **S**(8–5)**G**. Analogous to the MS/MS spectrum of **G**(*t*8–O–4)**G**, the cleavage of the β-aryl ether unit in **S**(*t*8–O–4)**S**(8–5)**G** was presumed to produce daughter ions corresponding with each fragment at *m/z* 225 and 387. Yet only the MS/MS ion (*m/z* 225) of the terminal 8–O–4-bonded **S** unit was clearly observed. The lack of a daughter ion associated with the dimeric 8–5-linked fragment was always observed in the case of phenylcoumaran-containing trimers (see Fig. Supp. Mat.). Presumably, the cleavage of the β-aryl ether unit results in the phenylcoumaran fragment with *m/z* 387 that rapidly eliminates water from the phenylcoumaran moiety leading to the product ion at *m/z* 369. Hence, **S**(8–O–4)**S**(8–5)**G** was proposed as structure for this compound, likely as its *threo*-isomer **S**(*t*8–O–4)**S**(8–5)**G**. This structure was further authenticated by the spiking and analysis of the synthesized compound.

*Compound **18***, **G**(*t*8–O–4)**S**(8–5)**G**

This compound had a molecular weight of 584 g/mol based on MS operated in the negative-ion mode and eluted at 16.1 min. The UV/VIS spectrum (λ_{max} = 226.5 and 274.7 nm) was similar to that of a phenylcoumaran. MS/MS daughter ions at *m/z* 535, 553 and 565, whereby the ion at *m/z* 535 was the base peak, were reminiscent of a β-aryl ether unit involving a **G** unit (*m/z* 195). As for the MS/MS spectrum of the previous compound, the pattern of peaks at *m/z* 354, 357 and 369 suggested again an **S**(8-5)**G** substructure. Hence, a **G** unit connected to **S**(8–5)**G** by means of an 8–O–4-bond was proposed. The *threo*-form of this compound, i.e. **G**(*t*8–O–4)**S**(8–5)**G** or buddlenol B (Houghton, 1985), was isolated from the $Cu(OAc)_2$ coupling reactions of coniferyl plus sinapyl alcohols and showed the same retention time, and UV/VIS and MS/MS spectra as those of the natural compound.

*Compound **19***, **SP**(8–8)**S**

At 16.4 min, a compound eluted with a UV/VIS spectrum characterized by λ_{max} at 205.4 and 236.5 nm. The molecular mass (556 g/mol) was in the range typically observed for trimers. MS/MS daughter ions at *m/z* 373, 387, 403 and 417 indicated a syringaresinol, i.e. **S**(8–8)**S** residue. Since *p*-hydroxybenzoic acids have been found acylating poplar lignins (Nakano et al., 1961; Meyermans et al., 2000), the neutral loss of 138 g/mol, generating the syringaresinol daughter ion at *m/z* 417, suggested a hydrated syringaresinol substructure carrying a γ-*p*-hydroxybenzoyl moiety. The compound could result from the 8–8-coupling of sinapyl *p*-hydroxybenzoate, i.e. an **SP** unit (see Materials and Methods), with sinapyl alcohol. This compound, **SP**(8–8)**S** (Lee et al., 1993) was synthesized (Lu et al., 2004) and showed an identical retention time, and UV/VIS and MS/MS spectra as the natural compound.

*Compound **20***, **G**(*e*8–O–4)**S**(8–5)**G**

An identical molecular weight (584 g/mol), UV/VIS (λ_{max} at 226.5 and 274.7 nm), and similar MS/MS spectra as for **G**(*t*8–O–4)**S**(8–5)**G** (compound **18**), isolated from synthetic $Cu(OAc)_2$ dehydrogenation, were obtained for the peak eluting at 16.9 min, suggesting that both compounds have similar molecular structures. The major difference in the MS/MS spectrum was the minor abundance of the product ion at *m/z* 535, which was the base peak in the spectrum of **G**(*t*8–O–4)**S**(8–5)**G**. This product ion, corresponding most likely with the loss of both H_2O and CH_2O from an 8-O-4-linked moiety, was still of the same magnitude as the peaks at *m/z* 553 (loss of CH_2O) and 565 (loss of H_2O). So, an *erythro*-configuration for the β-aryl ether unit is evident and **G**(*e*8–O–4)**S**(8–5)**G**, i.e. the *erythro*-diastereomer of buddlenol B (Houghton, 1985), was proposed as the most likely structure. Synthesized *threo*-buddlenol B (from the borate $Cu(OAc)_2$ oxidation of sinapyl plus coniferyl alcohols) contained minor amounts of the *erythro*-diastereomer, which eluted at the same retention time as compound **20**.

*Compound **21***, **S**(*t*8–O–4)**S**(8–5)**G'**

The UV/VIS spectrum (λ_{max} at 216.0, 278.3 and 344.7 nm) of the compound eluting at 17.8 min, showed absorption at longer wavelengths and suggested the presence of an aromatically conjugated aldehyde. The MS/MS spectrum indicated a molecular weight of 612 g/mol, 2 g/mol less than that of compound **17**, **S**(*t*8–O–4)**S**(8–5)**G**.

Many of the same neutral losses were observed in the MS/MS spectra of both compounds. The base peak at *m/z* 563, accompanied by peaks at *m/z* 581 and 593, indicated the presence of a β-aryl ether unit in the *threo*-configuration, whereas the relative abundances of the peaks at *m/z* 367, 355 and 352 were reminiscent of a phenylcoumaran unit containing a cinnamaldehyde group. The finding of a fragment ion at *m/z* 583 (neutral loss of 28 g/mol) further substantiated the presence of a carbonyl function. This compound therefore is **S**(*t*8–O–4)**S**(8–5)**G'**, the cinnamaldehyde analog of compound **17**.

*Compound **22***, **G**(*t*8–O–4)**S**(8–8)**G**

A compound with a similar UV/VIS spectrum (λ_{max} at 214.8 and 273.5 nm) to that of **S**(8–8)**S** (syringaresinol) eluted at 18.0 min. Its molecular weight of 584 Da, obtained by MS/MS analysis, indicated a trilignol composed of two **G** units and one **S** unit. The pattern of fragment ions at *m/z* 535 (base peak), 553 and 565 resulted from the dissociation of a *threo*-β-aryl ether. The cleavage of this linkage produced daughter ions at *m/z* 195 and 387. So, a **G** unit 8–O–4-linked to an **S**(8–8)**G** substructure, i.e. **G**(*t*8–O–4)**S**(8–8)**G** or buddlenol E (Houghton, 1985), was proposed. Spiking of the synthesized compound authenticated the structure.

*Compound **23***, **G**(*t*8–O–4)**S**(8–5)**G'**

The UV/VIS spectrum (λ_{max} at 228.8, 281.8 and 343.5 nm) of the peak eluting at 18.2 min, suggested the presence of a conjugated aldehyde. The molecular weight (582 g/mol) was two g/mol less than that of buddlenol B **18** and the same neutral losses were observed (See Fig. Supp. Mat.). Furthermore, the loss of 28 Da (*m/z* 553) indicated the presence of a carbonyl function. Hence, this product is the aldehyde analogue of buddlenol B **18**, i.e. buddlenol A **23** (Houghton, 1985). The relative abundances of the peaks at *m/z* 533, 551 and 563 in analogy to above, indicated that this compound was present in the *threo*-form, i.e. **G**(*t*8–O–4)**S**(8–5)**G'**. The structure was verified and confirmed by spiking of synthesized *threo*-buddlenol A **23**.

*Compound **24***, **G**(*t*8–O–4)**S**(8–8)**S**

This compound had a retention time of 18.3 min, a resinol-characteristic UV/VIS spectrum (λ_{max} at 213.6 and 279.4 nm) and a mass of 614 g/mol as determined by LC-MS/MS. The peak pattern composed of *m/z* 565, 583 and 595 suggested a *threo*-β-aryl ether unit involving a **G** unit (*m/z* 195) and a syringaresinol **S**(8–8)**S** **4** substructure (*m/z* 417). So, the **G**(*t*8–O–4)**S**(8-8)**S** structure, or buddlenol C (Houghton, 1985), was proposed and authenticated by spiking and MS/MS comparison with the synthesized compound.

*Compound **25***, **G**(*e*8–O–4)**S**(8–5)**G'**

At 18.8 min, a peak eluted having a molecular weight of 582 g/mol and the UV/VIS spectrum (λ_{max} at 231.2, 278.3 and 344.7 nm) suggested the presence of a conjugated aldehyde. Its MS/MS spectrum was almost identical to that of compound **23**, **G**(*t*8–O–4)**S**(8–5)**G'**. The main difference was the much lower abundant *m/z* 533, implying that this compound is the *erythro*-diastereomer of buddlenol A, i.e. **G**(*e*8–O–4)**S**(8–5)**G'**. Minute amounts of this isomer were present in synthesized *threo*-buddlenol A, which eluted at the same retention time and thus confirmed the proposed structure.

The structures of all following compounds but compound **33**, i.e. compounds **26** to **38**, were based on the interpretation of the MS/MS spectra in comparison with those of the structures described above.

*Compound **26***, **G**(8–5)**G**(8–O–4)**G'**

The separation on LC-MS/MS of fraction 16 (Fig. 3) revealed a compound with a molecular weight of 552 g/mol. The relative abundances of daughter ions at *m/z* 503, 521 and 533 were indicative of a *threo*-β-aryl ether unit. The MS/MS fragmentation produces the ions at *m/z* 177 and 373, which are presumably associated with a coniferaldehyde moiety and a **G**-type β-ether moiety. Thus, this compound is likely to be **G**(8–5)**G**(8–O–4)**G'**.

*Compound **27***, **G**(*t*8–O–4)**G**(8–5)**V'**

The MS/MS spectrum of this trilignol (MW = 544 g/mol), detected in fraction 17, comprised fragment ions at *m/z* 495 (base peak), 513 and 525, that could be attributed to a *threo*-β-aryl ether unit connecting a **G** unit (*m/z* 195) with a **G**(8–5)**V'** moiety (*m/z* 314, 317 and 329), confirming the structure as **G**(*t*8–O–4)**G**(8–5)**V'**.

*Compound **28***, **G**(*t*8–O–4)**G**(8–5)**G'**

LC-MS/MS analysis of fraction 20 revealed a trimer having a mass of 552 g/mol. The relative abundances of the daughter ions at *m/z* 503, 521 and 533 indicated the presence of a *threo*-β-aryl ether **G** unit (*m/z* 195). A phenylcoumaran substructure involving a **G** unit and a coniferaldehyde end group was indicated by peaks at *m/z* 337, 325 and 322. These data conform with **G**(*t*8–O–4)**G**(8–5)**G'**.

*Compound **29***, **G**(*e*8–O–4)**G**(8–5)**G'**

Fraction 21 contained a peak with a molecular weight of 552 g/mol based on LC-MS/MS. Fragmentation reactions occurring at an *erythro*-β-aryl ether unit yielded fragment ions at *m/z* 533, 521 and 503, whereas a phenylcoumaran unit involving an aromatic aldehyde could be inferred from the presence of daughter ions at *m/z* 337, 325 and 322. So, this compound is **G**(*e*8–O–4)**G**(8–5)**G'**, the *erythro*-isomer of compound **28**.

*Compound **30***, **S**(8–O–4*)**S**(8–5)**G**

A compound of mass 612 g/mol was present in fraction 22 as detected by LC-MS/MS. Daughter ions at *m/z* 563, 581 and 593 could be attributed to a β-aryl ether unit, whereas the presence of an **S**(8–5)**G** phenylcoumaran moiety was suggested by ions at *m/z* 354, 357 and 369. The ion at *m/z* 223 could be related to a 7-hydroxy-sinapaldehyde or a 7-oxo-sinapyl alcohol moiety involved in the β-aryl ether unit. However, the quinone methide formed during the 8–O–4-coupling of a sinapaldehyde is re-aromatized by loss of the 8–H with the formation of an α,β-enone structure rather than by the nucleophilic attack of a water molecule (Connors et al., 1970). Hence, only the presence of a 7–oxo-sinapyl alcohol residue, and thus an **S**(8–O–4*)**S**(8–5)**G** structure, fully accommodates the MS data and chemical coupling propensities.

*Compound **31***, **G**(*t*8–O–4)**S'**(8–8)**S** *or* **G**(*t*8–O–4)**S**(8–8)**S'**

LC-MS/MS analysis of fraction 27 showed a peak characterized by a molecular weight of 630 g/mol. MS/MS ion at *m/z* 581, 599 and 611 were reminiscent of a *threo*-β-aryl ether unit, whereby the ion at *m/z* 195 indicated the involvement of a **G** residue. The detection of ions at *m/z* 389, 403, 419 and 433 suggested an 8–8-linked dimer moiety. Both **G**(*t*8–O–4)**S'**(8–8)**S** and **G**(*t*8–O–4)**S**(8–8)**S'** are consistent with the MS/MS spectrum. The loss of 48 (H_2C=O and H_2O), 30 (H_2C=O) and 18 (H_2O) from the *m/z* 433 fragment concur with the acyclic nature of this moiety (as opposed to the tetrahydrofuran that could be formed by 9–OH attack on the quinone methide intermediate of the aldehyde moiety formed by 8–8-coupling of sinapaldehyde and sinapyl alcohol).

*Compound **32***, **G**(*e*8–O–4)**S**(8–8)**S**

In the same fraction, another peak was found with a mass of 614 g/mol and an MS/MS spectrum similar to **G**(*t*8–O–4)**S**(8–8)**S,** compound **24**. The near equal abundance of the daughter ions at *m/z* 565, 583 and 595 indicated an *erythro*-β-aryl ether unit instead of *threo*, and identified this compound as **G**(*e*8–O–4)**S**(8–8)**S** or *erythro*-buddlenol C.

Mixed-type tetramers

*Compound **33***, **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**

The peak eluting at 19.6 min had a UV/VIS spectrum (λ_{max} at 220.6 and 276.3 nm) that resembled the UV/VIS spectrum of a resinol. A molecular weight of 810 g/mol was obtained by LC-MS operated in the negative ion mode. This suggests a tetrameric structure composed of two **G** and two **S** units. MS/MS spectral peaks (See Fig.

Supp. Mat.) at *m/z* 791 (-H_2O), 773 (-2H_2O), 761 (-CH_2O and H_2O), 743 (-CH_2O and 2H_2O) and 713 (-2CH_2O and -2H_2O) indicated the presence of two β-aryl ether units. Further, the MS/MS daughter ion at *m/z* 613 fragmented into the following MS^3 granddaughter ions: *m/z* 565 (100%), 403 (80%), 417 (66%), 387 (34%), 583 (21%), 373 (18%) and 595 (17%). This MS^3 spectrum was similar to the MS/MS spectrum obtained for compound **24**, **G**(*t*8–O–4)**S**(8–8)**S**. As such, the structure **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G**, called hedyotisol (Matsuda et al., 1984), was proposed. The structure was confirmed by spiking with the authentic compound isolated from the $Cu(OAc)_2$ coupling reaction of sinapyl plus coniferyl alcohols. No *threo*/*erythro* configuration could be assigned to the 8–O–4-units based on the MS/MS spectrum.

*Compound **34***, **S**(*t*8–O–4)**SP**(8–8)**S** *or* **S**(*t*8–O–4)**S**(8–8)**SP**

LC-MS/MS analysis of fraction 22 (Fig. 3) revealed a compound having a molecular weight (782 g/mol) typically in the range of those observed for tetramers. The MS/MS spectrum showed the presence of a *threo*-β-aryl ether unit (*m/z* 733, 751 and 763). A neutral loss of 226 g/mol affording the daughter ion at *m/z* 555, suggested that this β-aryl ether unit linked an **S** unit to a trimeric moiety. The ion pattern at *m/z* 373, 387 and 417 indicated an **S**(8–8)**S** substructure. Since the ions at *m/z* 417 and 555 differ by 138 g/mol, a structure where an **S** unit is 8–O–4-coupled to **SP**(8–8)**S** (compound **19**), i.e. **S**(*t*8–O–4)**SP**(8–8)**S** or **S**(*t*8–O–4)**S**(8–8)**SP**, is evident. It is not possible from the MS data to determine whether the final sinapyl alcohol has added to the **SP** or the **S** end of the **SP**(8–8)**S** dimer **19**.

*Compound **35***, **G**(*e*8–O–4)**G**(8–O–4)**S'**(8–8)**S** *or* **G**(*e*8–O–4)**G**(8–O–4)**S**(8–8)**S'**

The presence of a compound with a mass of 826 g/mol was shown by LC-MS/MS of fraction 27. Daughter ions at *m/z* 777 and 807 suggested the presence of an *erythro*-β-aryl ether unit of which the MS/MS fragmentation leads to the ion at *m/z* 629. The presence of a second β-aryl ether unit involving a **G** unit is indicated by the differences between the ion at *m/z* 629 on one hand and the ions at *m/z* 611, 581 and 433 on the other. Due to the low abundance of the latter three ions, no distinction could be made between the diastereomers. The latter ions were further evidence for an **S**(8–8)**S'** moiety. Hence, two structures may be inferred from the MS/MS spectrum, i.e. **G**(*e*8–O–4)**G**(8–O–4)**S'**(8–8)**S** or **G**(*e*8–O–4)**G**(8–O–4)**S**(8–8)**S'**. Compound **35** logically derives from compound **31**, **G**(8–O–4)**S'**(8–8)**S** or **G**(8–O–4)**S**(8–8)**S'**, via coniferyl alcohol coupling.

*Compound **36***, **G**(8–O–4)**G**(8–8)**S**(8–O–4)**G**

In fraction 22, a compound was found with a molecular weight of 780 g/mol and an MS/MS spectrum that resembled that of hedyotisol, **G**(8–O–4)**S**(8–8)**S**(8–O–4)**G** **33**. Peaks at *m/z* 761, 743, 731, 713, 695 and 683 suggested the presence of a β-aryl ether unit at each end of this tetramer (see identification of hedyotisol **33**). However, the MS/MS data did not allow us to distinguish between the diastereomers. Peaks at *m/z* 613 (representing the trimeric fragment resulting from the cleavage of an 8–O–4-bond) and 417 (representing the syringaresinol **S**(8–8)**S** substructure remaining after cleavage of both 8–O–4-bonds) in the MS/MS spectrum of hedyotisol were downwardly shifted by 30 units toward *m/z* 583 and 387 in the spectrum of this compound, implicating **G**(8–O–4)**G**(8–8)**S**(8–O–4)**G** as the structure. No *erythro*/*threo* configuration could be assigned based on the MS/MS spectrum.

*Compound **37*** **S**(8–O–4)**S**(8–8)**S**(8–O–4)**G**

The MS/MS spectrum of this compound (MW = 840 g/mol) in fraction 25 showed the presence of a β-aryl linked unit at each end of the molecule, as suggested by the ions at *m/z* 821, 803, 791, 773, 755 and 743. Furthermore, the peak at *m/z* 417 could be associated with a syringaresinol substructure. Cleavage of the 8–O–4-bond on either side of the molecule produced two different daughter ions, at *m/z* 613 and 643, implying that the unit involved in one β-aryl unit is an **S**, whereas the other is a **G** unit. The *erythro*/*threo* configuration could not be derived from the MS/MS data.

*Compound **38***, **G**(*t*8–O–4)**G**(8–O–4)**S**(8–8)**S**

This compound (MW = 810 g/mol) was present in fraction 23. In the MS/MS spectrum, a *threo*-β-aryl ether

unit could be derived from the presence of peaks at *m/z* 761, 779 and 791. The peaks at *m/z* 417, and at *m/z* 613, 595 (*m/z* 613-H_2O), 583 (*m/z* 613-CH_2O) and 565 (*m/z* 613-H_2O-CH_2O) indicated the presence of a syringaresinol **S**(8–8)**S** substructure to which a **G** unit was attached via another 8–O–4-bond for which no *threo*/*erythro*-isomer assignment was possible.

Oligolignols not detected in xylem extracts

*Compound **39***, **S**(*t*8–O–4)**G**(8–5)**G**

This compound was synthesized via traditional organic chemistry methods (not involving radical coupling) to verify the correct interpretation of MS/MS spectra of phenylcoumaran-containing trilignols. This compound is not readily formed by *in vitro* dehydrogenative polymerization reactions because sinapyl alcohol only 8–O–4-cross-couples efficiently with **S** units (Syrjänen and Brunow, 1998). LC-MS/MS analysis of the isolated HPLC fractions of the xylem extracts did not reveal the presence of **S**(*t*8–O–4)**G**(8–5)**G**, in accordance with observed cross-coupling propensities. Nor was this compound detected in the synthetic **G**+**S** mixture (see below).

*Compound **40***, **G**(*t*8–O–4)**G**(8–5)**G**

This compound was detected in the reaction mixture from the $Cu(OAc)_2$ coupling reaction of sinapyl plus coniferyl alcohols (see below), but not in xylem extracts from poplar (Fig. 3).

LITERATURE CITED

Adler E (1977) Lignin chemistry — past, present and future. Wood Sci Technol **11:** 169-218

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol **54:** 519-546

Connors WJ, Chen C-L, Pew JC (1970) Enzymic dehydrogenation of the lignin model coniferaldehyde. J Org Chem **35:** 1920-1924

Cutillo F, D'Abrosca B, DellaGreca M, Fiorentino A, Zarrelli A (2003) Lignans and neolignans from *Brassica fruticulosa*: effects on seed germination and plant growth. J Agric Food Chem **51:** 6165-6172

Haruna M, Koube T, Ito K, Murata H (1982) Balanophonin, a new neo-lignan from *Balanophora japonica* Makino. Chem Pharm Bull **30:** 1525-1527

Houghton PJ (1985) Lignans and neolignans from *Buddleja davidii*. Phytochemistry **24:** 819-826.

Lee H, Watanabe N, Sasaya T, Ozawa S (1993) Extractives of short-rotation hardwood species I. Phenolics of the wood of *Salix sachalinensis* Fr. Schm. Mokuzai Gakkaishi **39:** 1409-1414

Lewis NG, Davin LB (1999) Lignans: biosynthesis and function. In U. Sankawa, ed, Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives, Comprehensive Natural Products Chemistry, Vol 1. Elsevier Science, Amsterdam, pp 639-712

Li Y-C, Kuo Y-H (2000) Four new compounds, ficusal, ficusesquilignan A, B, and ficusolide diacetate from the heartwood of *Ficus microcarpa*. Chem Pharm Bull **48:** 1862-1865

Lu F, Ralph J, Morreel K, Messens E, Boerjan W (2004) Preparation and relevance of a cross-coupling product between sinapyl alcohol and sinapyl *p*-hydroxybenzoate. Org Biomol Chem, in press

Matsuda S, Kadota S, Tai T, Kikuchi T (1984) Isolation and structures of hedyotisol-A, -B and –C, novel dilignans from *Hedyotis lawsoniae*. Chem Pharm Bull **32:** 5066-5070

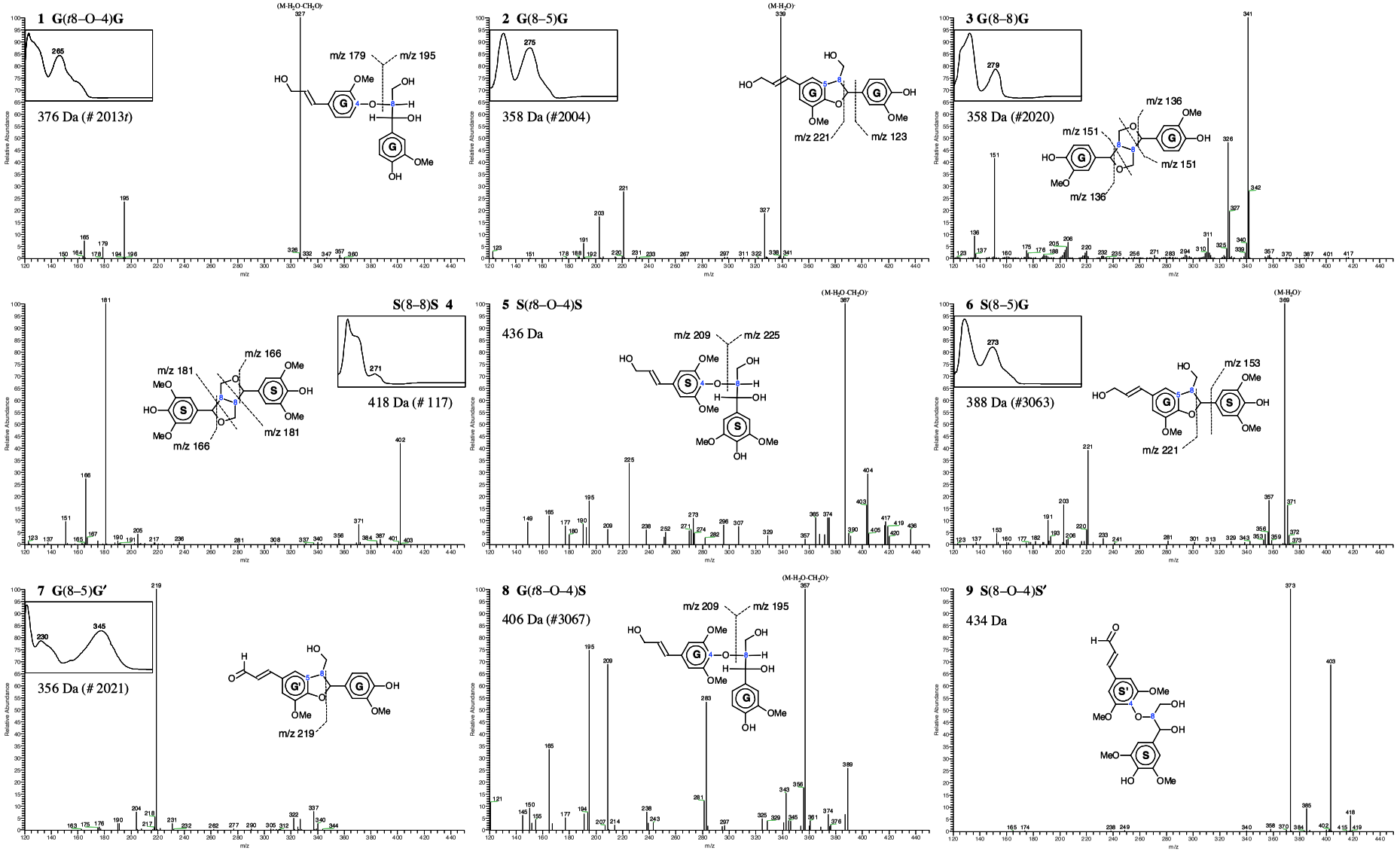
Meyermans H, Morreel K, Lapierre C, Pollet B, De Bruyn A, Busson R, Herdewijn P, Devreese B, Van Beeumen J, Marita JM, Ralph J, Chen C, Burggraeve B, Van Montagu M, Messens E, Boerjan W (2000) Modification in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. J Biol Chem **275:** 36899-36909

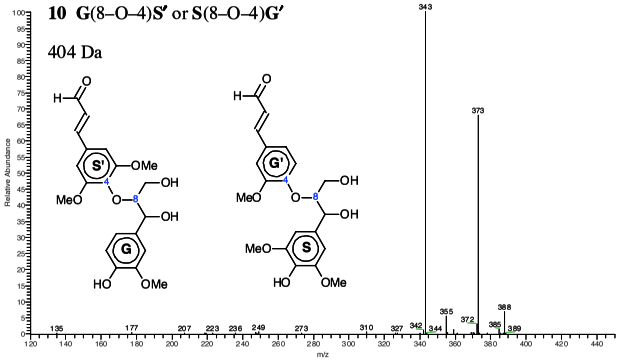
Nakano J, Ishizu A, Migata N (1961) Studies on lignin. XXXII. Ester groups of lignin. Tappi **44:** 30-32

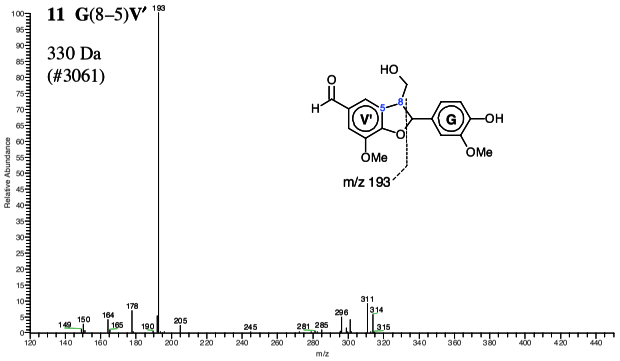
Syrjänen K, Brunow G (1998) Oxidative cross coupling of *p*-hydroxycinnamic alcohols with dimeric arylglycerol _-aryl ether lignin model compounds. The effect of oxidation potentials. J Chem Soc, Perkin Trans 1 **1998:** 3425-3429

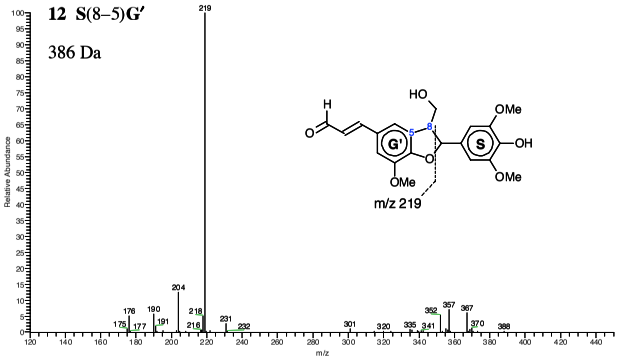
Umezawa T (2004) Diversity of lignan biosynthesis. Phytochem Rev, in press

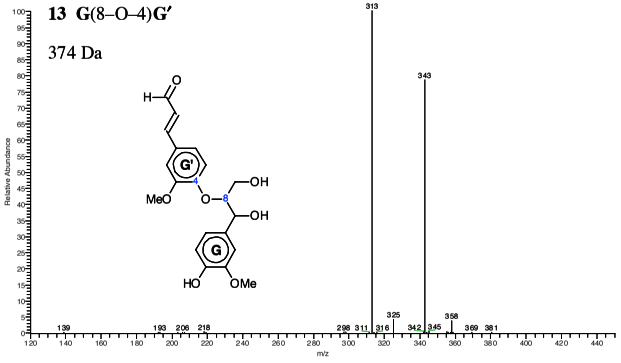
Yang Y-P, Cheng M-J, Teng C-M, Chang Y-L, Tsai I-L, Chen I-S (2002) Chemical and anti-platelet constituents from Formosan Zanthoxylum simulans. Phytochemistry **61:** 567-572

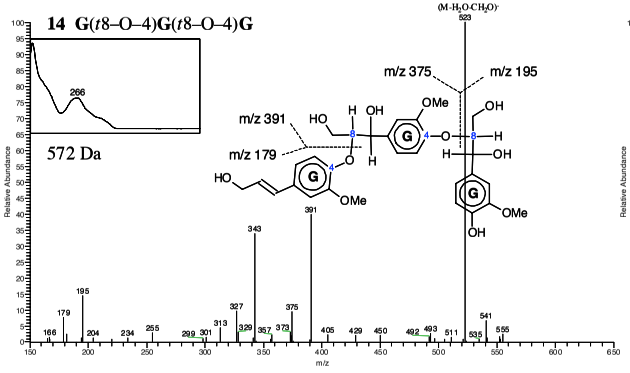

1 G(t8–O–4)G
376 Da (# 2013t)
m/z 179
m/z 195
2 G(8–5)G
358 Da (#2004)
m/z 221
m/z 123
3 G(8–8)G
358 Da (#2020)
m/z 136
m/z 151
S(8–8)S 4
418 Da (# 117)
m/z 166
m/z 181
5 S(t8–O–4)S
436 Da
m/z 209
m/z 225
6 S(8–5)G
388 Da (#3063)
m/z 153
m/z 221
7 G(8–5)G′
356 Da (# 2021)
m/z 219
8 G(t8–O–4)S
406 Da (#3067)
m/z 209
m/z 195
9 S(8–O–4)S′
434 Da
Relative Abundance
m/z

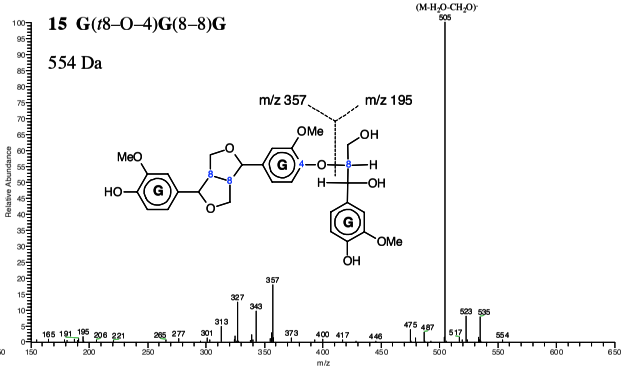

10 G(8–O–4)S′ or S(8–O–4)G′
404 Da
Relative Abundance
m/z

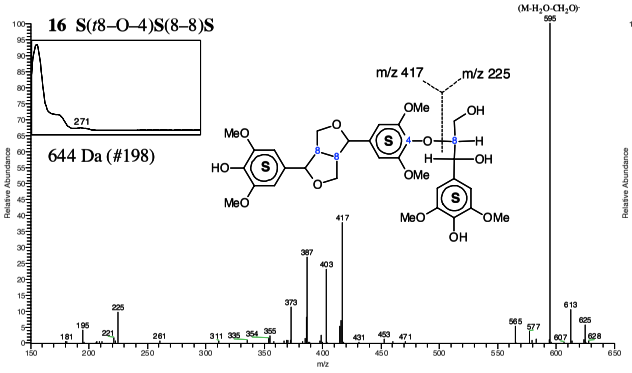

11 G(8–5)V′
330 Da
(#3061)
m/z 193
Relative Abundance
m/z

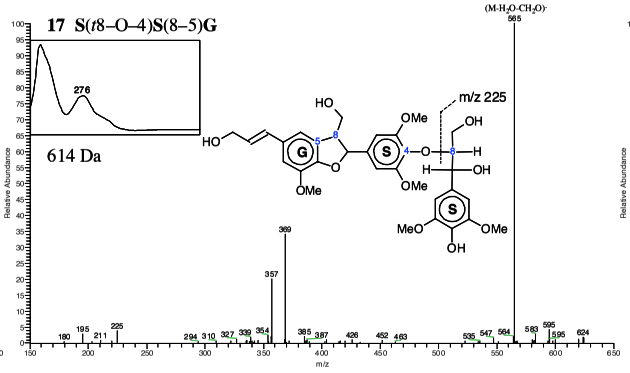

12 S(8–5)G′
386 Da
m/z 219
Relative Abundance
m/z

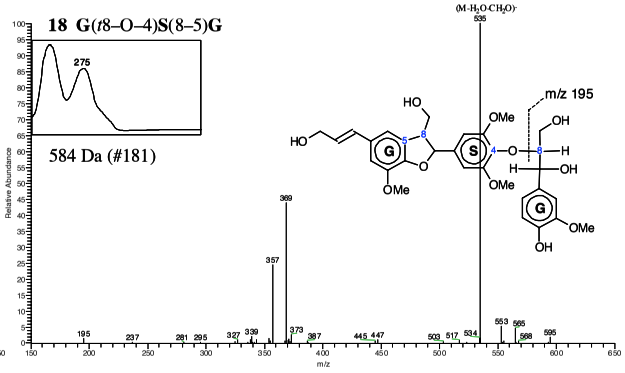

13 G(8–O–4)G′
374 Da
Relative Abundance
m/z

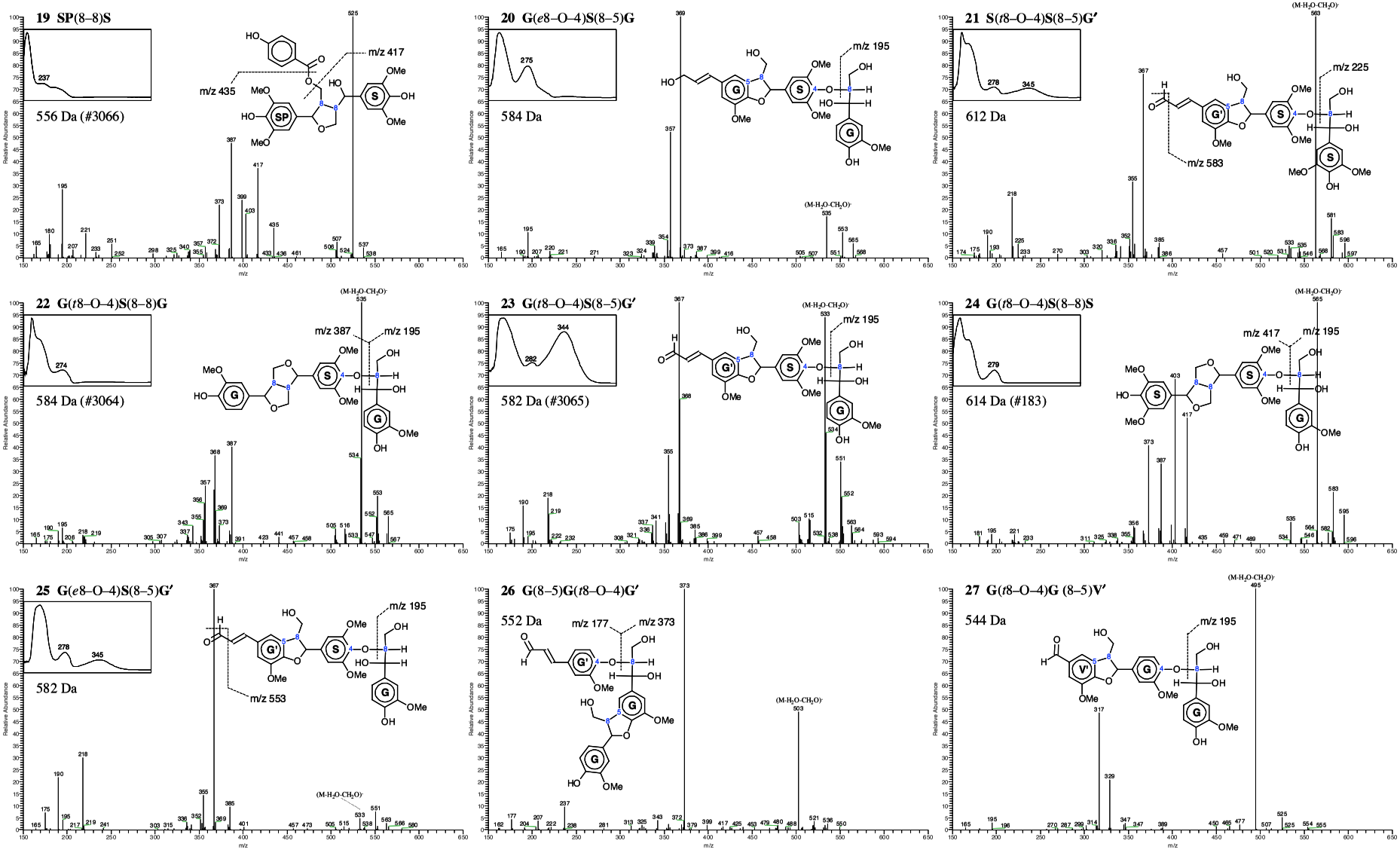

14 G(t8–O–4)G(t8–O–4)G
572 Da
m/z 375
m/z 195
m/z 391
m/z 179
Relative Abundance
m/z

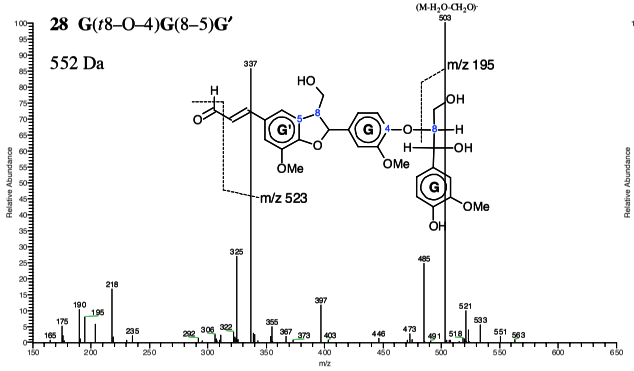

15 G(t8–O–4)G(8–8)G
554 Da
m/z 357
m/z 195
Relative Abundance
m/z

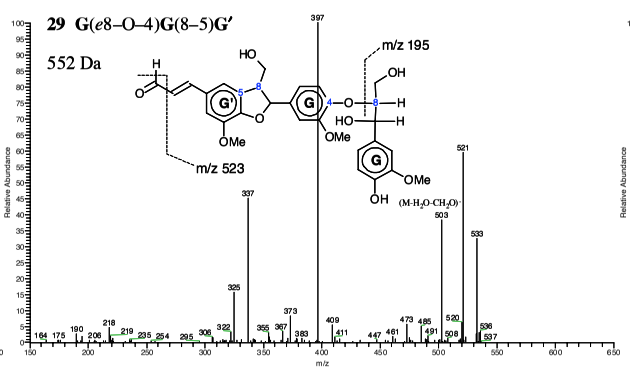

16 S(t8–O–4)S(8–8)S
644 Da (#198)
m/z 417
m/z 225
Relative Abundance
m/z

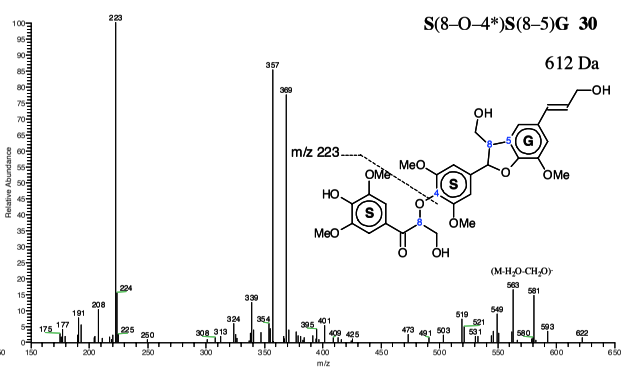

17 S(t8–O–4)S(8–5)G
614 Da
m/z 225
Relative Abundance
m/z

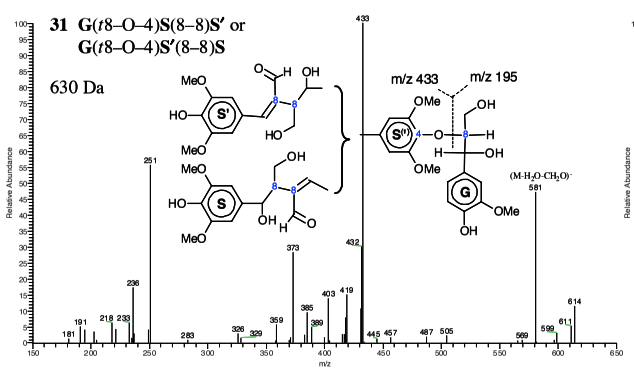

18 G(t8–O–4)S(8–5)G
584 Da (#181)
m/z 195
Relative Abundance
m/z

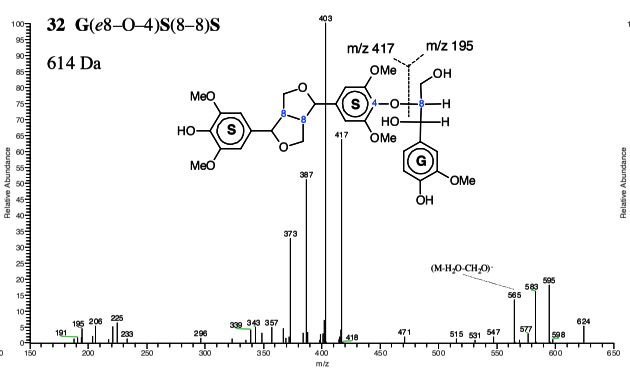

19 SP(8–8)S
556 Da (#3066)
m/z 417
m/z 435
20 G(e8–O–4)S(8–5)G
584 Da
m/z 195
21 S(t8–O–4)S(8–5)G′
612 Da
m/z 225
m/z 583
22 G(t8–O–4)S(8–8)G
584 Da (#3064)
m/z 387
m/z 195
23 G(t8–O–4)S(8–5)G′
582 Da (#3065)
m/z 195
24 G(t8–O–4)S(8–8)S
614 Da (#183)
m/z 417
m/z 195
25 G(e8–O–4)S(8–5)G′
582 Da
m/z 553
m/z 195
26 G(8–5)G(t8–O–4)G′
552 Da
m/z 177
m/z 373
27 G(t8–O–4)G (8–5)V′
544 Da
m/z 195
Relative Abundance
m/z

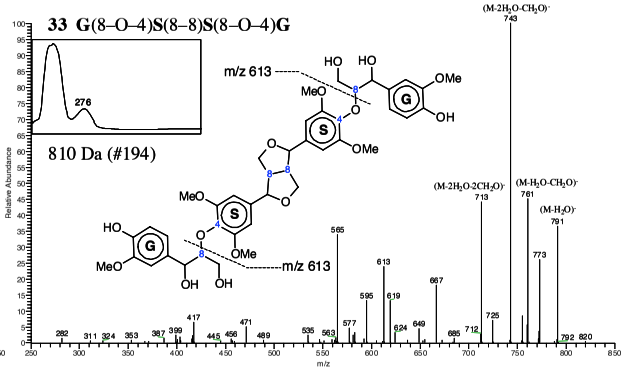

28 G(t8–O–4)G(8–5)G′
552 Da
m/z 523
m/z 195
Relative Abundance
m/z

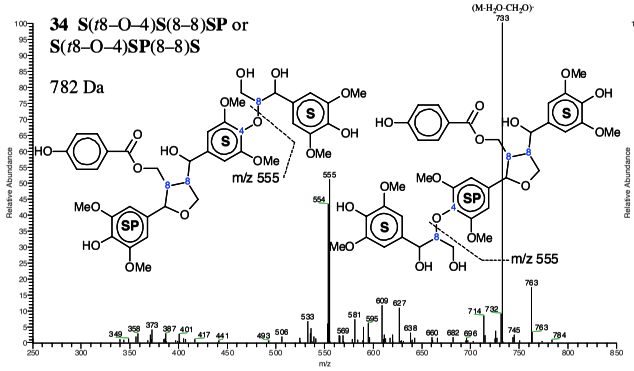

29 G(e8–O–4)G(8–5)G′
552 Da
m/z 523
m/z 195
Relative Abundance
m/z

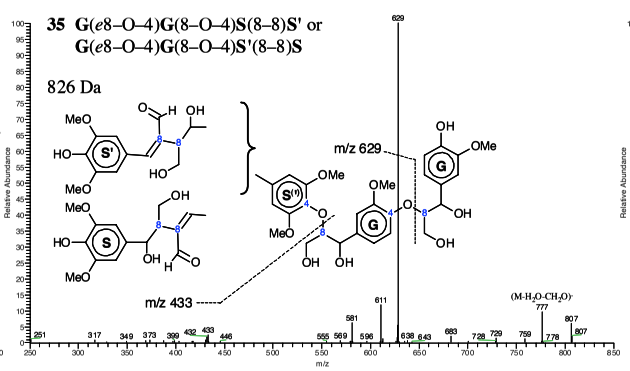

S(8–O–4*)S(8–5)G 30
612 Da
m/z 223
Relative Abundance
m/z

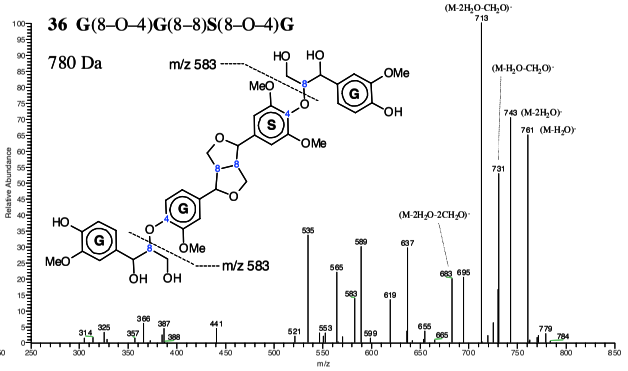

31 G(t8–O–4)S(8–8)S′ or
G(t8–O–4)S′(8–8)S
630 Da
m/z 433
m/z 195
Relative Abundance
m/z

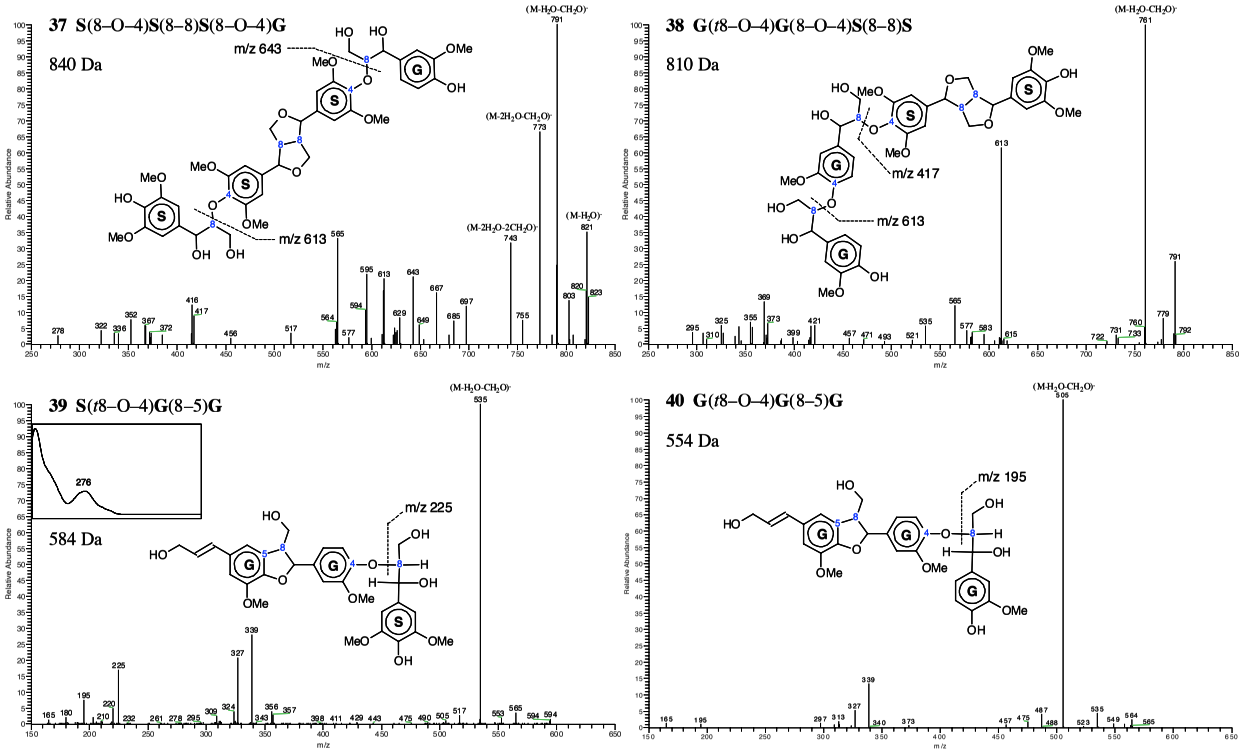

32 G(e8–O–4)S(8–8)S
614 Da
m/z 417
m/z 195
Relative Abundance
m/z


33 G(8–O–4)S(8–8)S(8–O–4)G
276
810 Da (#194)
m/z 613
Relative Abundance
m/z


34 S(t8–O–4)S(8–8)SP or
S(t8–O–4)SP(8–8)S
782 Da
m/z 555
Relative Abundance
m/z


35 G(e8–O–4)G(8–O–4)S(8–8)S' or
G(e8–O–4)G(8–O–4)S'(8–8)S
826 Da
m/z 629
m/z 433
Relative Abundance
m/z


36 G(8–O–4)G(8–8)S(8–O–4)G
780 Da
m/z 583
Relative Abundance
m/z


37 S(8–O–4)S(8–8)S(8–O–4)G
840 Da
m/z 643
m/z 613
(M-H2O-CH2O)-
791
(M-2H2O-CH2O)-
773
(M-2H2O-2CH2O)-
743
(M-H2O)-
821
38 G(8–O–4)G(8–O–4)S(8–8)S
810 Da
m/z 417
m/z 613
(M-H2O-CH2O)-
761
39 S(8–O–4)G(8–5)G
584 Da
276
m/z 225
(M-H2O-CH2O)-
535
40 G(8–O–4)G(8–5)G
554 Da
m/z 195
(M-H2O-CH2O)-
505
Relative Abundance
m/z